

(19)



(11)

EP 2 056 578 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
31.10.2012 Bulletin 2012/44

(51) Int Cl.:
H04M 3/51 ^(2006.01) **H04M 3/42** ^(2006.01)

(21) Application number: **08168071.2**

(22) Date of filing: **31.10.2008**

(54) **Providing a multi-modal communications infrastructure for automated call centre operation**

Bereitstellung einer multimodalen Kommunikationsstruktur für automatisierten Callcenterbetrieb

Fourniture d'une infrastructure de communications multimodale pour le fonctionnement automatisé d'un centre d'appel

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HR HU IE IS IT LI LT LU LV MC MT NL NO PL PT
RO SE SI SK TR**

(30) Priority: **31.10.2007 US 981860**

(43) Date of publication of application:
06.05.2009 Bulletin 2009/19

(73) Proprietor: **Intellisist, Inc.
Bellevue WA 98007 (US)**

(72) Inventors:
• **Odinak, Gilad
Bellevue, WA 98008 (US)**
• **Sutherland, Alastair
Seattle, WA 98103 (US)**

(74) Representative: **Curley, Donnacha John et al
Hanna Moore & Curley
13 Lower Lad Lane
Dublin 2 (IE)**

(56) References cited:
US-A1- 2005 147 090 US-A1- 2005 213 743

EP 2 056 578 B1

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

DescriptionTECHNICAL FIELD

5 **[0001]** The present invention relates in general to automated call center operation and, in particular, to a system and method for providing a multi-modal communications infrastructure for automated call center operation.

BACKGROUND ART

10 **[0002]** Customer call centers, or simply, "call centers," are usually the first direct point of contact for customers seeking direct assistance from manufacturers and service vendors. Call centers are reachable by telephone and provide a single source for customer support and problem resolution. Although World Wide Web-based customer support is becoming increasingly available via the Internet, call centers still offer a convenient and universally-available forum for remote customer assistance. In addition, for mobile callers, such as customers who are in a car and unable to access the Internet, call centers remain the only practical resource for receiving assistance on-the-go.

15 **[0003]** As customer satisfaction and good will depend significantly on service after the sale, vendors spend substantial time, money and effort in ensuring effective call center operation. Customer assistance professionals are trained in providing both effective and courteous communication and informed and accurate product and service information. Nevertheless, the volume of call traffic can often exceed the capabilities of human customer assistance agents, and a range of automated call center systems are presently used to help bridge the gap between the need to provide responsive assistance and the limits of human call center staff.

20 **[0004]** Typically, in existing automated call center systems, customers are put on hold until an agent is available to take their call. While on hold, an automated system typically collects information from the customer, such as account number, to determine a priority of service. Such a system may also provide automated menus that attempt to classify the call into basic transaction types, for instance, based on language spoken.

25 **[0005]** When an agent is available to take the call, the agent will greet the customer, may ask for identification information, and will attempt to determine the nature of the call, often by asking for the same information that the customer previously provided to the automated system. The agent then takes some action, such as performing troubleshooting and providing the caller with instructions for resolving the problem. Further, to assist in improving customer support, the agent will usually log the customer information provided by the caller.

30 **[0006]** Each customer call is typically an interactive process. During the interaction, the agent may put the customer on hold while the agent gathers information, takes notes, or sometimes handles other customers. Finally, after call completion, the agent will summarize the call as a call log entry. The interactive process is repeated for each new caller throughout the course of the day. This interactive process is time inefficient. Agents are forced to wait for customers to complete their instructions while customers similarly remain on hold while agents are assisting other callers, researching a problem resolution, or creating call logs.

35 **[0007]** A customer interaction system is described in U.S. Patent No. 6,778,660, to Fromm, issued on August 17, 2004, the disclosure of which is incorporated by reference. The system enables agents to simultaneously process voice contacts from telephonic callers by storing incoming voice signals for time-shiftable and fast playback. Calls are routed to assigned agents when possible. If an agent is busy, the user is asked to state his question for storage in one or more queues. An agent subsequently processes and responds to the question in person, in a manner similar to that utilized in Web chat interactions. Agents respond to voice recordings in the order in which received with additionally provided fast playback that enables an agent to catch up with recorded messages. However, both user and agent messages remain as spoken speech recorded in an audible, non-textual format and accordingly require the full attention of the assigned agent.

40 **[0008]** Accordingly, there is a need for an approach to providing automated call center operation that allows highly responsive caller support with a minimum of agent idle time and caller hold time.

45 **[0009]** There is a further need for an approach to providing efficient caller message processing using transcribed and synthesized speech utterances as an internal medium of communication within the automated call center.

50 **[0010]** US2005/147090 A1 discloses a multimedia call center with a client self-help system (Abstract). A transaction, such as a live call or media event, including email, is received by a communication center (paragraphs [0063] and [0088]). Once received, a type of the call or media event is determined and processing occurs based on the type (paragraphs [0064] and [0088]). The content of the transaction received at the communication center is recorded and stored according to enterprise rules (paragraphs [0070] and [0089]). Information is extracted from the transaction for building a knowledge base that can be used for dynamic management and future analysis (paragraph [0090]).

55 **[0011]** US2005/213743 A1 discloses transferring a customer to a live agent when predetermined criteria are met (Abstract). A customer first communicates with a conversational agent via a text query (paragraph [0025]). The conversational agent is a computer program that evaluates the query provided by the customer (paragraph [0026]). If unable

to resolve the query, the conversational agent escalates the customer to the live agent via one of a telephone call, instant message, and mobile text message (Id.). The customer can be escalated to the live agent through a link, an email address or telephone numbers, or by joining the current conversation (paragraph [0086]).

5 DISCLOSURE OF THE INVENTION

[0012] In the described embodiment, an agent and customer communicate through voice messages using a digitized voice-driven system. From the customer perspective, the experience appears to be an interaction with an intelligent machine. The interaction is similar to calling a legacy automated call center system staffed with human agents, but the customers are aware that the agent is automated, not human.

[0013] Preferably, the system voice is clear and human-like, but is recognizable as a machine voice. Slight delays in responses can occur compared to speaking with a human agent, although the customer does not have to repeat information already provided and is generally not put on hold. Upon request, the system can repeat the information provided to the customer, and consistently appears to be patient and polite.

[0014] Operationally, the system differs from legacy systems. Instead of providing full-time voice-to-voice communications, the system gives agents the ability to control a continuum of increasingly automated responses in the form of a "sliding" control. For most interactions, every customer speech utterance is recorded and stored, is digitally transcribed into a text message and is presented to an off-line agent through a visual interface on a workstation. The agent can read or optionally listen to each utterance upon arrival, as well as to previous utterances. As well, the agent can annotate or manually re-transcribe each of the customer utterances as needed.

[0015] Once a session has been established, each human agent can communicate indirectly with customers by typing written responses at their workstation. Each written response is converted into speech following completion and is played to the customer.

[0016] The agent can also choose pre-formed responses, thereby saving time and communicating in a consistent, uniform manner. The pre-formed responses can include an associated form containing parameterized variable fields that are completed by the agent or by the system to fill in, for example, dates or names. The completed pre-formed response is converted into speech. Alternatively, pre-formed responses can be pre-recorded as sound bites and staged in an audio cache for immediate playback.

[0017] In addition to indirect communication via their workstation, each human agent can also accept live calls from customers in a real time, voice-to-voice mode. Live call processing may be required to handle crises or other service needs that are not amenable to automation, or to provide those services to customers, which are not generally acceptable when provided via automation.

[0018] Communication can also occur between a customer and a human agent via text messaging, such as through the Short Message Service (SMS), as well as multi-modal communication, which includes a combination of voice and text messaging. The text communications can be automatically reviewed and stored as a log entry without requiring transcription. Multi-modal communication allows a customer to text information to the agent during the call session for providing additional data.

[0019] Furthermore, when an agent identifies a call that matches a recognized problem or frequently asked question, the agent can choose a predefined "script" to prompt and collect or simply provide the customer with information in a step-by-step manner. For example, a script could be used to collect a customer's personal information for a credit application, or to provide instructions on using a feature of a consumer product in a customer support application. Thus, the ability of an agent to interact with customers through manually or automated text responses converted into speech or through pre-recorded or live voice responses provides a flexible and sliding level of agent control adaptable to a wide range of customer service situations.

[0020] The system also provides an automatic journaling function. By the time each call ends, the system will have collected a complete and fully transcribed log of the conversation. The human agent need not manually transcribe a log entry, as the information contained in the system-generated log is already in a format that can be stored in a database and can be easily mined for data. Manual annotations can be added to the log entry, as needed.

[0021] Empirically, from an agent perspective, an average customer service call lasts seven minutes when using a legacy call center, of which two minutes are spent on an introduction and setup and an additional two minutes on wrap-up and documentation. The described embodiment eliminates most of the time spent on setup and wrap-up and reduces the time an agent spends on a call by about 30%. From a customer point of view, although the overall length of the call may be the same, hold times are eliminated while slight yet unobjectionably longer delays between agent's responses may occur.

[0022] With the extensive logging and visual presentation to the agent, the agent can keep track of more than one session with a customer. An agent may handle multiple calls simultaneously, while each customer will hardly notice any degradation in call quality. In the described embodiment, an agent can handle four calls simultaneously. However, an arbitrarily large number of simultaneous calls could be maintained by the system for each agent and would only be

subject to physical computational constraints, such as available processing power and memory. Each agent is able to handle multiple calls by not having to listen to each customer as they speak. Instead, agents can listen to or read transcribed customer utterances and text messages, thereby each allowing their attention to be split among multiple calls. The ability to use pre-recorded forms to gather or provide standard information further increases an agent's ability

to handle multiple calls. For example, while a customer is interacting with a script collecting personal information, such as first and last name, social security number, address and phone number, and so forth, the agent can handle other callers. **[0023]** An embodiment is a system and method for providing a message-based communications infrastructure for automated call center operation is described. A call from a user into a call center is accepted. The accepted call includes a stream of transcribed verbal speech utterances. Each transcribed verbal speech utterance is recorded as a user message. The accepted call is assigned to a session, which is then assigned to an agent. The call is progressively processed in the assigned session by presenting each user message to the assigned agent, executing commands responsive to the assigned agent, and sending an agent message to the user. The agent message includes a stream of synthesized audible speech utterances.

[0024] A further embodiment is a system and method for efficiently operating an automated call center through text-based messaging. A voice-based caller is interfaced through a telephonic medium. Audible speech utterances are exchanged during a call into an automated call center. Each call is managed as a session and each session is assigned to an agent. Incoming audible speech utterances are converted into machine-processable text messages provided to the agent assigned to the session. An action specified by the assigned agent is executed. Outgoing machine-processable text messages are processed into audible speech utterances provided to the voice-based caller transacting the session.

[0025] A further embodiment is a system and method for providing a multi-modal communications infrastructure for automated call center operation. A multi-modal call is accepted from a caller through a telephony interface, which accommodates multi-modal calls including at least one of verbal speech and text messaging. Incoming speech in the multi-modal call is converted into transcribed text. Incoming text messaging is matched with the transcribed text. The multi-modal call is automatically assigned through a session manager to a session under operation of a live agent. The transcribed text and incoming text messaging are progressively processed during the session through an agent application by performing a customer support scenario interactively monitored and controlled by the live agent.

[0026] In yet another embodiment a system is described for providing a multi-modal communications infrastructure for automated call center operation. This system comprises; a telephony interface configured to accept a multi-modal call from a caller, which accommodates multi-modal calls comprising at least one of verbal speech and text messaging; a speech recognition module configured to convert incoming speech in the multi-modal call into transcribed text; a text message processor to match incoming text messaging with the transcribed text; a messaging server configured to automatically assign the multi-modal call through a session manager to a session under operation of a live agent; and an agent console configured to progressively process the transcribed text and incoming text messaging during the session through an agent application by performing a customer support scenario interactively monitored and controlled by the live agent.

[0027] In a further embodiment a method is provided to facilitate a multi-modal communications infrastructure for automated call center operation. The method comprises the steps of accepting a multi-modal call from a caller through a telephony interface, which accommodates multi-modal calls comprising at least one of verbal speech and text messaging; converting incoming speech in the multi-modal call into transcribed text; matching incoming text messaging with the transcribed text; automatically assigning the multi-modal call through a session manager to a session under operation of a live agent; and progressively processing the transcribed text and incoming text messaging during the session through an agent application by performing a customer support scenario interactively monitored and controlled by the live agent.

[0028] In yet another embodiment a system for providing a multi-modal communications infrastructure for automated call center operation is described. The system comprises a telephony interface configured to accept a multi-modal call from a caller, which accommodates multi-modal calls comprising at least one of verbal speech and text messaging; a messaging server configured to automatically assign the multi-modal call through a session manager to a session under operation of a live agent; a speech recognition module configured to convert incoming speech in the multi-modal call into transcribed text; a text message processor configured to match incoming text messaging with the transcribed text; and a messaging server configured to automatically assign the multi-modal call through a session manager to a session under operation of a live agent; and an agent console configured to progressively process the transcribed text and the incoming text messaging during the session through an agent application by performing a customer support scenario interactively monitored and controlled by the live agent.

[0029] The system may further comprise a database configured to store the incoming text messaging as a log entry and to match the log entry with a stored copy of the transcribed text. The incoming text messaging and the incoming speech may simultaneously originate from the caller. The system may further comprise a caller identification module configured to assign a caller identification to the incoming text messaging; and the messaging server configured to forward the incoming text messaging to the live agent based on the caller identification. The system may further comprise a post-processing module configured to perform post-processing on the incoming text messaging. The system may

further comprise an identification module configured to flag the incoming text messaging as verbatim caller data; and a database configured to store the flagged incoming text messaging. The live agent may process the incoming speech by listening to the multi-modal call and by performing an action comprising at least one of annotating and manually transcribing the incoming speech.

[0030] In yet another embodiment, A method is described for providing a multi-modal communications infrastructure for automated call center operation, comprising: accepting a multi-modal call from a caller through a telephony interface, which accommodates multi-modal calls comprising at least one of verbal speech and text messaging and automatically assigning the multi-modal call through a session manager to a session under operation of a live agent; converting incoming speech in the multi-modal call into transcribed text; matching incoming text messaging with the transcribed text; and automatically assigning the multi-modal call through a session manager to a session under operation of a live agent; and progressively processing the transcribed text and the incoming text messaging during the session through an agent application by performing a customer support scenario interactively monitored and controlled by the live agent.

[0031] The method may further comprise storing the incoming text messaging as a log entry; and matching the log entry with a stored copy of the transcribed text. The incoming text messaging and the incoming speech simultaneously originate from the caller. The method may further comprise: assigning a caller identification to the incoming text messaging; and forwarding the text messaging to the live agent based on the caller identification. The method may further comprise performing post-processing on the incoming text messaging. The method may further comprise: flagging the incoming text messaging as verbatim caller data; and storing the flagged incoming text messaging. The method may further comprise processing, by the live agent, the incoming speech by listening to the multi-modal call; and performing an action comprising at least one of annotating and manually transcribing the incoming speech.

[0032] In an example a system is detailed for providing service provisioning to a caller through multi-modal communication. The system comprises a message server configured to assign an incoming multi-modal call from a caller to a call session having a unique caller identification and which is managed by an agent; a telephony interface configured to receive verbal speech communication from the caller during the call session; a speech recognition module configured to convert the verbal speech to transcribed text and further configured to provide the transcribed text to the agent; a text gateway configured to receive incoming text messages from the caller; an identification module to assign the unique caller identification to each incoming text message and further configured to provide the incoming text messages to the agent based on the unique caller identification; and a text-to-speech engine configured to receive outgoing text messages from the agent and further configured to convert the outgoing text messages into an outgoing stream of synthesized speech that is presented to the caller.

[0033] The verbal speech communication and the incoming text messages of the embodiments simultaneously originate from the caller. The system of the example may further comprise: a communication interface configured to receive the verbal speech communication and the incoming text messages into a call center through separate streams of communication and further configured to pair the verbal speech communication and the incoming text messages, comprising: a caller identification module configured to assign the unique caller identification to the verbal speech communication; and a call match module configured to match the unique caller identification of the verbal speech communication and the incoming text messages to the caller session. The system may further comprise: a database configured to store the transcribed text and the incoming text messages to a call log. The system may also further comprise: a caller information module configured to receive the multi-modal call from a mobile communication device, further configured to obtain caller information from the mobile communication device comprising locational data, and further configured to analyze the locational data to determine a location of the caller.

[0034] Another example is for providing service provisioning to a caller through multi-modal communication. This method comprises: assigning an incoming multi-modal call from a caller to a call session having a unique caller identification and which is managed by an agent; and providing service provisioning to the caller during the multi-modal call, comprising: receiving verbal speech communication from the caller during the call session and converting the verbal speech to transcribed text; providing the transcribed text to the agent; receiving incoming text messages from the caller and assigning the unique caller identification to each incoming text message; providing the incoming text messages to the agent based on the unique caller identification; and receiving outgoing text messages from the agent and converting the outgoing text messages into an outgoing stream of synthesized speech that is presented to the caller. The verbal speech communication and the incoming text messages of the embodiments simultaneously originate from the caller. The method of the example may further comprise receiving the verbal speech communication and the incoming text messages into a call center through separate streams of communication; and pairing the verbal speech communication and the incoming text messages, comprising: assigning the unique caller identification to the verbal speech communication; and matching the unique caller identification of the verbal speech communication and the incoming text messages to the caller session. The method may further comprise storing the transcribed text and the incoming text messages to a call log. The method may further comprise receiving the multi-modal call from a mobile communication device; obtaining caller information from the mobile communication device comprising locational data; and analyzing the locational data to determine a location of the caller.

[0035] Still other embodiments of the present invention will become readily apparent to those skilled in the art from the following detailed description, wherein is described embodiments of the invention by way of illustrating the best mode contemplated for carrying out the invention. As will be realized, the invention is capable of other and different embodiments and its several details are capable of modifications in various obvious respects, all without departing from the spirit and the scope of the present invention. Accordingly, the drawings and detailed description are to be regarded as illustrative in nature and not as restrictive.

DESCRIPTION OF THE DRAWINGS

[0036]

FIGURE 1 is a functional block diagram showing an automated call center operational environment.

FIGURE 2 is a process flow diagram showing, by way of example, a typical user call sequence, as transacted in the automated call center operational environment of FIGURE 1.

FIGURE 3A is a block diagram showing a system for providing a message-based communications infrastructure for automated call center operation, in accordance with one embodiment.

FIGURE 3B is a block diagram showing a system for providing a message-based communications infrastructure for automated call center operation, in accordance with a further embodiment.

FIGURE 4 is a process flow diagram showing variable automation levels provided using the system of FIGURE 3A.

FIGURE 5 is a block diagram showing a schema for storing records in the logging database of the system of FIGURE 3A.

FIGURE 6 is a process flow diagram showing, by way of example, the control flow followed in processing a script using the system of FIGURE 3A.

FIGURE 7A is a screen shot showing, by way of example, a set of call center service windows generated by the system of FIGURE 1.

FIGURE 7B is a screen shot showing, by way of example, a single call center service window with multiple call tabs, generated by the system of FIGURE 1.

FIGURE 8 is a flow diagram showing a method for providing a message-based communications infrastructure for automated call center operation, in accordance with one embodiment.

FIGURE 9 is a flow diagram showing a routine for assigning a session for use in the method of FIGURE 8.

FIGURE 10 is a flow diagram showing a routine for processing a session for use in the method of FIGURE 8.

FIGURE 11 is a flow diagram showing a routine for processing a session by an agent for use in the routine of FIGURE 10.

FIGURE 12 is a flow diagram showing a routine for processing a session by a script engine for use in the routine of FIGURE 11.

FIGURE 13 is a flow diagram showing a routine for executing a script for use in the routine of FIGURE 12.

FIGURE 14 is a block diagram showing a mobile messaging system for presenting output data, in accordance with a further embodiment.

FIGURE 15 is a flow chart diagram showing a routine for presenting output data using the mobile messaging system of FIGURE 14.

FIGURE 16 is a flow chart diagram showing a routine for displaying presentation content based on a vehicle condition.

FIGURE 17 is a block diagram showing a vehicle system for communicating with the system of FIGURE 3A.

FIGURE 18 is a flow chart diagram showing a method for communicating with the system of FIGURE 3A using the vehicle system of FIGURE 17.

APPENDIX A provides a sample grammar for use in the described embodiment.

BEST MODE FOR CARRYING OUT THE INVENTION

[0037] FIGURE 1 is a functional block diagram showing an automated call center operational environment 10. By way of example, a multiplicity of users call into an automated call center 11, preferably through telephonic means, which can include voice or text messaging media. The telephonic means include Plain Old Telephone Service (POTS) 12, cellular and satellite telephones 13, Internet telephony (IPTel) 14, Voice over IP (VoIP) 15, short message service (SMS) text messaging, and other forms of telephony and voice-based communications, as would be recognized by one skilled in the art. Users could also call or interface to the automated call center 11 through data transmission means, such as an internetwork 17, including the Internet. The call can include voice communications, text messaging, or a combination of voice and text communications.

[0038] Independent of call origination, each user call is routed through a telephone company (Telco) public interchange

16 or equivalent call center to the automated call center 11. Although shown with reference to a Telco public interchange 16, any other form of telephonic or equivalent call networking system transmitting voice or data signals over various signal carrier mediums, including conventional land lines; radio, satellite or other forms of signal carriers; light wave or sound wave exchange systems; or equivalents thereof, could also be utilized, as would be recognized by one skilled in the art.

[0039] The automated call center 11 provides a single source for support and problem resolution for customers seeking direct assistance from manufacturers and service vendors, although automated call centers 11 can also be used in other areas of commerce, as would be recognized by one skilled in the art. The terms "user" and "customer" are used interchangeably herein and both refer to a caller to the automated call center 11. Although the automated call center 11 is shown as a single point within the automated call center operation environment 10, the automated call center 11 could consist of one or more logically interconnected but physically separate, including geographically removed, operations, which provide a logically unified automated call center, as would be recognized by one skilled in the art.

[0040] FIGURE 2 is a process flow diagram showing, by way of example, a typical user call sequence 20, as transacted in the automated call center operational environment 10 of FIGURE 1. Upon calling into the automated call center 11, each user receives an initial greeting and an informational message providing a synopsis of caller options. The caller options enable the user to navigate through to a specific topic area for assistance or support. Following application, the user engages in a customer support scenario 22 with an agent, which is either a live person or an automated prompt, such as with an automated voice response system, to enable information collection and problem trouble-shooting. Note that the scenario 22 can be delayed by agent unavailability, caller volume capacity limits, and other factors that can delay customer response. As required, service provisioning 23 is provided to the user, either directly in the course of the call or indirectly through a service request dispatch. Finally, the call ends in a wrap-up 24, which provides closure to the call and a departing salutation. Other forms and variations of customer call sequences are feasible, as would be recognized by one skilled in the art. Importantly, however, from the perspective of the caller, the experience appears to be an interaction with an intelligent machine and the caller is aware that the agent is automated, not human. Accordingly, a typical caller will have a more relaxed expectation of agent responsiveness since a machine, and not an actual person, is on the line.

[0041] FIGURE 3A is a block diagram showing a system 30 for providing a message-based communications infrastructure for automated call center 11 operation, also referred to as the Messaging Platform for Agent-Customer Transactions (MPACT) system 30, in accordance with the present invention. During regular operation, the MPACT system 30 executes multiple threads to process multiple simultaneous calls, which are handled by one or more agents executing agent applications 43 on agent consoles 39 (shown in FIGURE 1). Alternatively, in a further embodiment, multiple MPACT systems 30 execute in parallel to provide enhanced performance through loosely- or tightly-coupled parallel processing.

[0042] The MPACT system 30 consists of the following components: database 34, telephony interface (TI) 35, one or more speech recognition engines (SREs) 36, one or more text-to-speech (TTS) engines 37, audio cache 38, one or more agent consoles 39, and optionally, resource manager 40. At least one instance of each component is generally required for the MPACT system 11 to operate, except that use of the resource manager 40 is optional, and is required only on larger systems that incorporate more than one instance of the other components.

[0043] In the described embodiment, a messaging server 31, database 34, telephony interface 35, SREs 36, TTS engines 37 and audio cache 38 execute on a single computer system while one or more agent consoles 39 executing in parallel on separate computer systems. The different components communicate over an Internet Protocol (IP) network, which typically is implemented over high-speed local Ethernet. The MPACT system 30 components run on Intel/AMD-based servers under the Windows 2000 Server Operating System and Redhat Linux. Each agent console 39 runs on Intel/AMD-based workstations under the Windows 2000 Professional Operating System. Each of the components will now be described in further detail.

[0044] The individual computer systems, including MPACT system 30, are general purpose, programmed digital computing devices consisting of a central processing unit (CPU), random access memory (RAM), non-volatile secondary storage, such as a hard drive, network interfaces, and peripheral devices, including user interfacing means, such as a keyboard and display. Program code, including software programs, and data are loaded into the RAM for execution and processing by the CPU and results are generated for display, output, transmittal, or storage.

[0045] Each component is implemented as is a computer program, procedure or module written as source code in a conventional programming language, such as the C++ programming language, and is presented for execution by the CPU as object or byte code, as is known in the art. Alternatively, the components could also be implemented in hardware, either as integrated circuitry or burned into read-only memory components. The various implementations of the source code and object and byte codes can be held on a computer-readable storage medium or embodied on a transmission medium in a carrier wave. The MPACT system 30 operates in accordance with a sequence of process steps, as further described below with reference to FIGURE 8.

[0046] The messaging server 31 consists of a script engine 32 and session manager 47. The script engine 32 executes scripts 49 incorporating a grammar 33 that defines a set of executable instructions for specified and interactive question-

and-response dialog, and a session manager 47 that includes a message queue 48 for staging transient user and agent messages. Script pointers ("Ptrs") 28 that identify the actual scripts 49 to be executed by the script engine 32 are maintained locally by each agent application 43. Alternatively, scripts 29 could be stored in the database 34 and the script pointers 28 would provide database pointers to the scripts 29. The messaging server 31 receives call control information from the telephony interface 35 and tracks logged-off, logged-on and signed-on agents. The messaging server 31 uses this information to establish sessions between agents and customers, as further described below with reference to FIGURE 4.

[0047] The database 34 contains three primary specialized databases: log database (LDB) 44, agent profile database (APDB) 45, and agent application database (AADB) 46. The log database 44 provides a running journal as a structured log of each accepted call and stores each user message and agent message in a uniquely-identified record. Thus, agents are freed from the task of having to manually transcribe a telephone conversation following wrap-up. The agent profile database 45 allows the messaging server 31 to authenticate, log-on and sign-on agents into registered sessions. The agent application database 46 contains agent applications 43 that are executed on agent consoles 39. Optionally, the database 34 can also contain scripts 29. Other forms of specialized databases are possible, as would be recognized by one skilled in the art. Alternatively, the information stored in the log database 44, agent profile database 45, and agent application database 46 could be maintained in structured or unstructured form using a standard file, spreadsheet, or other data assemblage for information storage and retrieval, as is known in the art.

[0048] Definitionally, a *signed-on agent* is registered on the messaging server 31 and is actively processing calls. A *logged-on agent* is registered on the messaging server 31 but is not accepting calls. A *logged-off agent* is not registered on the messaging server 31.

[0049] For each session, the messaging server 31 receives customer calls through the telephony interface 35 and sends a stream of transcribed speech utterances as user messages to an agent assigned to handle the session. Note that one or more agents can be assigned to handle any given session and a hierarchy of areas of responsibility, such as speech transcription, customer interaction, controlling scripts, and so forth, can be delegated among several agents to ensure efficient call processing. Similarly, the messaging server 31 receives a stream of synthesized speech utterances as agent messages from an assigned agent application 43 and sends the agent messages to the customer through the telephony interface 35. The messages typically only contain digitized voice; however, Simultaneous Voice and Data (SVD), for example, Caller ID, can also be provided. The multiplexing and demultiplexing of SVD messages is handled at the telephony interface 35 and an agent console 39.

[0050] The script engine 32 executes individual scripts 49, which incorporate a pre-defined grammar 33. The grammar 33 specifies a set of instructions that can be used to create question-and-answer dialogs that are executable by an agent via the agent application 43 and thereby enables an agent to process simultaneously multiple calls. The scripts 49 are submitted by agents via an agent console 39 using the script pointers 28 during processing of customer calls, as further described below with reference to FIGURE 6. Each script 49 defines a sequence of synthesized speech utterances sent to customers and transcribed speech responses received back as user messages. The speech utterances could be pre-recorded and staged in the audio cache 38. The user messages are used to populate a form (not shown) that is reviewed by the agent during service provisioning.

[0051] In a further embodiment, a call session can be multi-modal, that is, consisting of voice, text messaging, or a combination of both. The telephony interface 35 implements a text messaging gateway, in addition to providing conventional POTS and voice-based telephony services. The telephony interface 35 enables the messaging server 31 to accommodate both voice communications and text messaging, which can be voice- or text-only, or a combination of voice and text during any single session. Customer calls received by the call center through the telephony interface 35 are forwarded to the messaging server 31, which handles their processing according to type. For instance, text messages can be forwarded directly to an agent, whereas voice communications can require voice recognition processing before being displayable. Although both text messages and voice communications simultaneously originate from the same customer, each might travel along a different path to the call center, depending upon the underlying telephone service. However, through a text message processor (not shown), the messaging server 31 can pair up separate streams of text messages and voice communications matching caller identification data to an on-going caller session.

[0052] In one example text messages are received through the Telco via an SMS gateway (not shown) implemented in the telephony interface 35, or other component, such as a data network gateway (not shown) that is directly interfaced to the internetwork 17 or other public data communications network. Upon receipt by the call center, the text messages are transmitted to the messaging server 31 for processing. A unique caller identification for call session identification is assigned to each text message. The messaging server can match each text message to an active customer call. For instance, based on shared caller identification data, the guide to which an active call is assigned can be located by the database for forwarding the text message to that guide. The text messages are also added to the database after the call session has terminated as part of the call log. Post-processing actions can be performed on saved text messages in a manner similar to voice communications. In addition, text messages can be specifically flagged as verbatim caller data, which would not be unreliable due to, for instance, improper voice recognition processing.

[0053] In a still further example, communication between a customer and an agent at a call center can occur solely through text messaging. The text messages are similarly identified using unique caller identification data and the text messages received during a single call session can be stored together.

[0054] In a still further example, a mobile caller, such as a driver of a vehicle, can interact with a call center as further described with reference to FIGURES 14-18. The mobile communications can include voice communication, text messaging, or a combination of voice and text communication. The messaging server 31 is able to retrieve mobile caller information in the same fashion as any other caller. However, additional data about the caller, which is unique to mobile use, may also be provided, such as locational data through Global Positioning Satellite interface.

[0055] FIGURE 3B is a block diagram showing a system 160 for providing a message-based communications infrastructure for automated call center operation, in accordance with a further embodiment. A scripting server 161 executes as a separate system from the messaging server 31, which preferably includes only the session manager 47 and message queue 48. Providing the functionality of the script engine 162 on a scripting server 161 enhances overall system throughput and performance by delegating script processing on a system separate from the messaging server 31.

[0056] The scripting server 161 consists of a dedicated script engine 162, which executes scripts 163 stored locally to the scripting engine 161. The scripts 163 also incorporate the grammar 33. Script pointers ("Ptrs") 164 that identify the actual scripts 163 to be executed by the script engine 162 are maintained locally by each agent application 43. Alternatively, scripts 29 (shown in FIGURE 3A) could be stored in the database 34 and the script pointers 164 would provide database pointers to the scripts 29.

[0057] Referring back to FIGURE 3A, customer calls are received through the telephony interface 35, which provides the external connection between the MPACT system 30 and the telephone company 16 (shown in FIGURE 1). The primary purpose of the telephony interface 35 is to accept and process conventional telephone signals, including multiplexing, call routing, and queueing, as is known in the art. In the described embodiment, the telephony interface 35 consists of a third party hardware interface and software drivers, plus MPACT proprietary software that connects the third party package to the messaging server 31 and, in large systems, also to the resource manager 40. The MPACT system 30 supports standard telephony interface cards, such as analog and T1 Dialogic PCI cards. Optionally, the telephony interface 35 includes an audio cache 41 in which pre-recorded "canned" sound bites are stored for efficient playback. These sound bites provide informational and navigational messages to all callers. Optionally, two or more telephony interfaces 35 can be used to provide increased user call capacity. Other forms of telephony interface 35 could be used to accommodate various signal carrier mediums, including conventional land lines; radio, satellite or other forms of signal carriers; light wave or sound wave exchange systems; or equivalents thereof, as would be recognized by ones skilled in the art.

[0058] User calls consist of ordinary spoken words, which must be transcribed into written text, for display, processing and storage. The purpose of the speech recognition engine 36 is to generate a stream of transcribed speech utterances that are recorded as computer-processable user messages. In the described embodiment, the speech recognition engine 36 consists of third party software and MPACT proprietary software that connects the third party package to the agent application 43 and, in large systems, also to the resource manager 40. The MPACT system 30 supports Speechworks and Nuance speech recognition engines. Optionally, two or more speech recognition engines 36 can be used to provide increased user call capacity.

[0059] In a further embodiment, a speech recognition engine executes on a client system 18 interfaced to the MPACT system 30 over the internetwork 17, or other data transmission means. The MPACT system 30 receives client messages already transcribed into text by the client system 18 for processing in the same manner as live calls received directly through the telephony interface 35.

[0060] Each caller into the MPACT system 30 receives feedback in the form of agent messages, which each consist of a stream of synthesized speech utterances. The synthesized speech utterances include greetings, questions, informational responses, and other spoken phrases that the user hears during the course of the session. The purpose of the text-to-speech engine 37 is to generate the stream of synthesized speech utterances formed as agent messages, which can be played back as reproducible audio. The text-to-speech engine 37 consists of third party software and MPACT proprietary software that connects the third party package to the agent application 43 and, in large systems, also to the resource manager 40. MPACT system 30 supports Speechworks' Speechify text-to-speech. Optionally, the text-to-speech engine 37 can include an audio cache 42 that stores pre-recorded "canned" sound bites, which provide efficient delivery of standardized synthesized speech utterances for use in scripts and repetitive agent actions. As well, two or more text-to-speech engines 37 can be used to provide increased user call capacity.

[0061] In a further embodiment, a text-to-speech engine executes on the client system 18 interfaced to the MPACT system 30 over the internetwork 17, or other data transmission means. The MPACT system 30 sends agent messages to the client system 18 for synthesis into speech. The synthesized speech is heard by the caller on the client system 18 in the same manner as a live call through telephonic means transacted directly through the telephony interface 35.

[0062] Each agent console 39 provides the primary means for direct customer interaction. The primary purpose of each agent console 39 is to execute one or more agent applications 43, stored in the agent application database 46,

which display both user and agent messages and provide menus of actions that can be executed responsive to agent commands, including script execution, as further described below with reference to FIGURE 6. In the described embodiment, one or more agent applications 43 execute on each agent console 39 and one or more agent consoles 39 execute in parallel. Alternatively, multiple instances of agent applications 43 can run on a server machine (not shown) and can be accessed by agents at agent consoles 39 operating as remote terminals.

[0063] Each agent application 43 implements a graphical user interface (GUI) for the human agent. FIGURE 7A is a screen shot showing, by way of example, a set of call center service windows 91, 92, 93 generated by the system 10 of FIGURE 1. Each call center service window 91, 92, 93 appears in a graphical user interface 90 and enables an agent to indirectly interact with a customer calling through the telephony interface 35. Following sign-on, via the agent application 43, an agent can accept new session requests from the messaging server 31 and create a visual session container for each session.

[0064] In the described embodiment, up to four sessions can be presented to an agent simultaneously. Preferably, the agent can view the contents of all sessions on a single screen. One session is designated as the active session and accepts agent commands, such as an instruction to listen to a transcribed user message, play a synthesized agent message to the customer, or activate a script through a menu 94, as further described below with reference to FIGURE 6. The agent can switch active sessions with a single keystroke or pointer click.

[0065] The multiple call sessions can also be displayed one at a time in the user graphical interface. FIGURE 7B is a screen shot showing, by way of example, a single call center service window 95 with multiple call tabs, generated by the system 10 of FIGURE 1. The call center service window 95 is displayed in a graphical user interface to enable an agent to communicate with a customer. Multiple call tabs 96 are located near the top of the window 95 to indicate multiple active call sessions. Each call session can be identified by a corresponding customer name, phone number, or identification number, as well as by other identifiers. The call session identifiers can be displayed in the call tabs. Each call center service window 95 also includes caller identification fields 97, a caller profile 98, and a reason for the call 99. To navigate to a different call center service window from the window 95 displayed, the agent can select the appropriate call tab 96 corresponding to the call session.

[0066] Referring back to FIGURE 3A, each agent application 43 receives a stream of transcribed speech utterances from the telephony interface 35 via the messaging server 31. In an alternate embodiment (not shown), the stream of transcribed speech utterances bypasses the messaging server 31 and is received directly from the telephony interface 35. The messaging server 31 communicates with the speech recognition engine 36 to transcribe the utterances before sending the transcribed utterances to the agent application 43. In turn, the agent application 43 sends agent messages to the telephony interface 35 via the messaging server 31. The messaging server 31 communicates with the text-to-speech engine 37 to convert an agent message into a stream of synthesized speech utterances prior to forwarding to the telephony interface 35.

[0067] FIGURE 4 is a process flow diagram showing variable automation levels 50 provided using the system 30 of FIGURE 3A. A typical caller sequence proceeds in three stages: input 51, processing 52, and output 53. During the processing stage 52, the system 30 provides each agent with a "sliding" control 64 that can vary the level of automation used in customer service provisioning. At one end of the sliding control 64, the agent must manually type each written response to a user inquiry and, at the other end, the system 30 directly carries out customer interactions in a fully-automated fashion.

[0068] The sliding control 64 accommodates the need to provide linear and non-linear processing to flexibly and dynamically tailor call transaction processing. Non-linear processing occurs when a caller departs from an expected course of action, such as by providing a tangential response to a question in a script requesting specific information. The agent would read the user message and alter the course of script processing to accommodate the tangential response, thereby enhancing the comfort level of the customer towards problem resolution. Linear processing occurs when the system 30 interacts directly with the caller through pre-defined scripts and is able to successfully complete a series of steps towards problem resolution along a pre-defined course of action. During a course of a call, both linear and non-linear processing can be used strategically to increase user confidence level and to sufficiently process a larger volume of caller traffic than possible with a fully manual and fully non-linear call processing approach.

[0069] During the input stage 51, speech 54 is received into the automated call center 11 (shown in FIGURE 1), either by a human operator 55 or a machine 56, such as the MPACT system 30. The speech 54 is converted into a stream of transcribed speech utterances or "normal" text annotations 57. The text annotations 57 are machine-processable as inputs to the processing stage 52. A message manager (not shown) associates incoming the audible speech utterances with the text annotations 57 and the associated incoming audible speech utterances are provided with the text annotations 57 to the agent for optional playback.

[0070] Processing generally involves the execution of some form of action, such as a script execution, as further described below with reference to FIGURE 5. During the processing stage 52, either a human operator 58, that is, an agent, or a machine 59 interprets each annotation 57 and generates a response or query 60. The response or query 60 is received in the output stage 53, either by a human operator 61, that is, an agent, or a machine 62. Finally, the response

or query 60 is formed into synthesized speech utterances 63 that are played back to the caller.

[0071] In the described embodiment, the three stages of call processing, input 51, processing 52, and output 53, are performed by machines 56, 59, 62, although adjunctive processing can be performed by human operators 55, 58, 61, as necessary to supplement the MPACT system 30. Thus, a sliding control 64 of human operator and machine interaction can be provided to automate call center operations. Using the sliding control 64, the agent can change the behavior of the script engine 32 (shown in FIGURE 3A) by beginning execution of a script 29, adjusting the point of execution within a script 29, or by causing a portion of the script 29 to be repeated. The agent can thereby alter the ordinarily sequential control flow of script processing by intervening as necessary, based on the current context of the call, or can allow script processing to proceed in a linear fashion.

[0072] In addition to indirect communication via an agent console 39, each agent can also accept live calls from customers directly through the telephony interface 35 in a real time, voice-to-voice mode. Live call processing may be required to handle crises or other service needs that are not amenable to automation, or to provide those services to customers, which are not generally acceptable when provided via automation.

[0073] In a further embodiment, the agent communicates with callers executing on client systems 18 through text message exchange transacted over the internetwork 17, or other data transmission means. Unlike conventional chat sessions, caller processing is transacted in the same manner in which telephonic calls received through the telephony interface 35 are transacted. Consequently, the agent can apply the sliding control 64 over automation level to ensure a continuum of non-linear to linear call processing independent of the actual call transmission means. Various arrangements and combinations of call transmission means can therefore be accommodated, as would be recognized by one skilled in the art.

[0074] Referring back to FIGURE 3A, as a first step, an agent, via an agent console 39, authenticates to the messaging server 31 prior to logging on. Following log-on, the agent indicates availability to handling calls by signing on to the messaging server 31. Thereafter, calls that have been accepted into the automated call center 11 (shown in FIGURE 1) are assigned to a session, and the session is subsequently assigned to a signed-on agent.

[0075] After the initial communication between the agent and the customer, an agent application 43 ordinarily loads a script describing the session flow from the application database 45, either according to agent instructions or based on information provided by the telephony interface 35, such as Caller ID. The agent application 43 communicates further via the messaging server 31 with the speech recognition engine 36 and text-to-speech engine 37 to transcribe text-to-speech and convert speech-to-text, as necessary. Communication between the agent application 49 and the speech recognition engine 36 and text-to-speech engine 37 continues until the call terminates.

[0076] FIGURE 5 is a block diagram showing a schema 70 for storing records in the log database 44 of the MPACT system 30 of FIGURE 3A. Each session is identified by a session identifier (ID) 71, which uniquely identifies each session, such as a sequentially-increasing number. In addition, each record contains a message identifier (ID) 72, time stamped 73, sender or receiver flag 74, and content 75, typically constituting audio data, text data or notes transcribed by the agent. Other types of identifiers, formation and content can be stored in a record, as would be recognized by one skilled in the art.

[0077] Referring back to FIGURE 3A, each agent application 43 also contains a supervisor mode that can be used to manually monitor system behavior and performance, control agent application 43 and messaging server 31 operation, monitor and guide human agents actions, and perform similar administrative tasks. A separate administrator application (not shown) enables system setup and configuration.

[0078] Finally, each agent application 43 can execute scripts 49 to perform a pre-defined sequence of instructions, generally consisting of question-and-response dialogues through which a customer is queried regarding a concern or to troubleshoot a problem. FIGURE 6 is a process flow diagram showing, by way of example, the control flow 80 followed in processing a script 49 using the system 30 of FIGURE 3A. During the initial stage of processing, a customer 82 indirectly interacts with an agent 81 through the execution of an initial script 83. The purpose of the initial script 83 is to populate a standard form 84 with general information regarding the nature of the call. Upon reviewing the form 84, the agent 81 executes, via the agent application 43, one or more additional scripts 87 to provide problem resolution or troubleshooting and to receive further information via additional forms 88. Finally, the agent, via the agent application 43, generates a resolution form 86 that is used as parameters to a closing script 85, which is executed for playback to the customer 82 to complete the call.

[0079] In the described embodiment, each form 84, 86, 88 is structured as a data containment object that stores data relevant to the agent application 43. Preferably, each data containment object is maintained as a binary large object (BLOB) interpretable by each agent application 43 based on business requirements. Significantly, the use of scripts, forms and agent interaction enables a non-linear execution path through problem resolution and troubleshooting. As necessary, an agent, through the agent application 43, can manually enter data into a form and progressively modify the sequence of problem resolution and troubleshooting. The amount of manual agent intervention follows from the sliding control 64 implemented in the MPACT system 30, as described above with reference to FIGURE 4.

[0080] Referring back to FIGURE 3A, the script engine 32 executes each script 49, which incorporate the grammar

33. By way of example, a sample grammar for use in the described embodiment is provided in Appendix A. Other forms of grammars and scripting languages could be used, as would be recognized by one skilled in the art.

[0081] The resource manager 40 provides scalability, load balancing and redundancy in large systems comprising multiple speech recognition engines 36, text-to-speech engines 37, and telephony interfaces 35. In the described embodiment, the messaging server 31 has a built-in simple resource manager 40 (not shown) to manage multiple agent applications 43 operating in smaller call capacity systems.

[0082] Using an administration application that controls the resource manager 40, an administrator can set and configure the system while operational. The resource manager 40 enables the administrator to add or remove servers and to reroute connections between different components, for instance, between telephony interface 35, messaging server 31 and agent application 43.

[0083] The audio cache 38 provides a centralized repository in which pre-recorded "canned" sound bites are stored for efficient playback. These sound bites provide both informational and navigational messages to all callers and standardized synthesized speech utterances for use in scripts and repetitive agent actions. The sound bites in the audio cache 38 are retrievable by the telephony interface 35, text-to-speech engine 37 and agent application 43.

[0084] In a typical system, different components run on separate machines. A typical medium-sized system consists of one server running a telephony interface 35 and messaging server 31, a separate server for the speech recognition engine 36, another server for the text-to-speech engine 37, and a fourth server for the log, agent profile, and agent application databases 44, 45, 46, respectively. Alternatively, a minimal system runs all the components on a single server, along with an agent application 43 on an integrated agent console 39.

[0085] In the described embodiment, each medium-sized system configuration is "24×6," meaning that the system can handle 24 simultaneous calls and can interface with six human agents. A minimal system configuration is "4×1," that is, four simultaneous calls with one human agent, while a large system configuration is "96×24," that is, 96 simultaneous calls and 24 human agents. Through the resource manager 40, an aggregation of the above-described configurations enables much larger call capacity systems.

[0086] By mixing proprietary technologies and existing systems, the MPACT system 30 reduces the average call time by 30%, increases agent efficiency up to four-fold, and significantly improves customer satisfaction.

[0087] FIGURE 8 is a flow diagram showing a method for providing a message-based communications infrastructure 100 for automated call center operation, in accordance with the present invention. The method is executed by the MPACT system 30 and individual operations are executed by the various components, specifically described below. During regular operation, the MPACT system 30 processes multiple simultaneous calls, which are handled by one or more agents executing agent applications 43 on an agent console 39 (shown in FIGURE 1).

[0088] Generally, the method 100 proceeds by iteratively processing each call in a continuous processing cycle. During each cycle, a call is received (block 101) and assigned to a session (block 102) by the session manager 47 (shown in FIGURE 3A), as further described below with reference to FIGURE 9. Next, the session is processed (block 103), as further described below with reference to FIGURE 10. Following session processing, the call ends (block 104) and further call are processed (block 105) until the MPACT system 30 is shut down or no further calls are received. The method then terminates. In the described embodiment, the MPACT is a multi-threaded system, employing multiple threads, which each independently execute the method 100.

[0089] FIGURE 9 is a flow diagram showing a routine for signing a session 110 for use in the method 100 of FIGURE 8. The purpose of the routine is to instantiate a new session object for processing by an agent application 43 (shown in FIGURE 3A) and to subsequently assign the instantiated session to an agent operating on an agent console 39.

[0090] Initially, if any agent is available (block 111), one of the agents is picked as the assigned agent (block 112) and the new session is created (block 113). Subsequently, the selected agent is assigned to the newly-created session (block 114). The routine then returns. If no agents are available (block 111), the customer is presented with the option of interfacing to a non-agent (block 115), that is, an automated voice response system, which provides the information specifically requested by the customer (block 116), after which the routine returns. Otherwise, if the customer prefers an agent (block 115), the customer enters into a waiting queue (block 117) until an agent becomes available.

[0091] Note that both the customers and agents can be prioritized using predefined selection criteria. For instance, customers who have enrolled in premium support service can received a higher priority in the waiting queue than other customers. As well, specialized problem-resolution agents can be prioritized for servicing particular customer needs for consideration during selection of agent assignment.

[0092] FIGURE 10 is a flow diagram showing a routine 125 for processing a session 120 for use in the method 100 of FIGURE 8. The purpose of the routine is to iteratively store and forward messages using the message queue 48 (shown in FIGURE 3A). Other forms of queueing and message handling and prioritization are possible, as would be recognized by one skilled in the art.

[0093] During each iteration, the session manger 47 (shown in FIGURE 3A) waits for a message, either a user message or agent message (block 121). Upon receipt, the message is placed in the message queue 48 (block 122) and the recipient of the message is notified (block 123). If the message is a user message being sent to an agent (block 124),

the message is processed by the agent assigned to the session to which the user message corresponds (block 125), as further described below with reference to FIGURE 11. Iterative processing continues with each subsequent message (block 126), after which the routine returns.

[0094] FIGURE 11 is a flow diagram showing a routine for processing a session by an agent 130 for use in the routine 125 of FIGURE 10. The purpose of the routine is to facilitate the interaction between an agent and customer through an agent application executing on an agent console 39 (shown in FIGURE 3A).

[0095] First, the notification message is displayed (block 131) on the graphical user interface 90 (shown in FIGURE 7) of the agent application 43. As necessary, the agent sends agent messages to the customer from service provider or script engine 32 (shown in FIGURE 3A) to provide appropriate handling of the user message (block 132). If the sent message consists of a request to execute a script 49 (block 132), the message is further processed by the script engine 32 (block 134), as further described below with reference to FIGURE 12. The routine then returns.

[0096] FIGURE 12 is a flow diagram showing a routine for processing a session by a script engine 140 for use in the routine 130 of FIGURE 11. The purpose of this routine is to iteratively process each script execution request using the script engine 32 on behalf of requesting agent applications 43 (shown in FIGURE 3A).

[0097] Each message thread is iteratively processed (blocks 141-146) as follows. During each iteration (block 141), the notes field of each message is retrieved (block 142) and, if a script execution request is found (block 143), a script 49 is executed (block 144), as further described below with reference to FIGURE 13. Otherwise, if no script request is present (block 143), no operation occurs (block 145). Processing continues with each additional message thread (block 146), after which the routine returns.

[0098] Although described above with reference to the linear processing of a script in sequential order, agent intervention in script processing is fully supported. The agent continues to monitor the progress of the script execution by observing user responses and can intervene as necessary to accommodate a non-scripted response. For example, the user may provide a tangential response to a question in the script requesting specific information. The agent would read the user message and alter the course of script processing to accommodate the tangential response to the sliding control 64 (shown in FIGURE 4).

[0099] FIGURE 13 is a flow diagram showing a routine for executing a script 150 for use in the routine 140 of FIGURE 12. The purpose of this routine is to perform standard retrieve-and-interpret script instruction execution, as is known in the art.

[0100] First, each instruction is retrieved (block 151) and executed (block 152). In the described embodiment, instruction execution follows from an interpretable stored grammar 33 (shown in FIGURE 3A) and as provided, by way of example, in Appendix A. If the executing instruction is a "Quit" instruction (block 153), the script execution terminates and the routine returns. Otherwise, script processing continues with the next instruction (block 154).

[0101] The automated call center can also be helpful to mobile users to provide directions, event information, and assistance, as well as to perform concierge and other services for the mobile user. FIGURE 14 is a block diagram 160 showing a mobile messaging system for presenting output data, in accordance with a further embodiment. The mobile messaging system (MMS) 161 can include a user interface 162 having a microphone 164 for capturing the user's voice; a display 165; speakers 166; and an interface entry 167 for allowing the user to perform various interactive functions. The MMS 161 also includes a processor 168; a global positioning system (GPS) 169 for determining precise vehicle locations; and a communications device 163, such as a cellular modem, for transmitting and receiving wireless information. The MMS 161 is also coupled to receive information from other vehicle systems (not shown), such as a speedometer, an engine, or a braking system.

[0102] FIGURE 15 is a flow chart diagram showing a method for presenting output data using the system of FIGURE 14. The processor 168 of the MMS 161 receives presentation content (block 171). The presentation content is preferably received by the communications device 163 from a server transmitted over a network (not shown). The processor 168 determines if any vehicle emergency condition exists for a vehicle (block 172) based on information received from vehicle systems other than the MMS. If the processor 168 determines that a vehicle emergency condition exists, the processor 168 stops all output of received presentation content (block 173) to any one of the components of the user interface 162. Examples of a vehicle emergency condition that can cause the processor to stop output of the presentation content to the user interface 162 are signals received from the engine that indicate engine overheating or a signal that indicates the driver of the vehicle has applied brakes to the point where an anti-lock braking system is activated. After the output is stopped, the processor 168 can again receive presentation content (block 171).

[0103] If the processor 168 determines that there is not a vehicle emergency condition, the processor 168 receives vehicle condition information (block 174). The vehicle condition information comes from a number of sources that generate signals, such as a GPS 169 or speedometer. The processor 168 outputs the presentation content through the user interface 162 based on the vehicle condition information (block 175). The substance and format of the presentation content displayed on the display 165 or broadcast on the speakers system 166 of the user interface 162 is determined. For example, when the vehicle is parked, the processor 138 presents all or nearly all of the received presentation content. As the vehicle condition changes, for example, by increasing speed, the processor 138 changes the presentation of the

content information shown on the display 165.

[0104] The change in presentation content can also involve shifting the presentation of the information from a visual display to an audio broadcast, as output through the speakers 166. In a further example, the presented content includes a full detailed map of the vehicle's location with directions to a previously requested destination. Based on the vehicle speed, or the rate of change of the vehicle speed or direction, the detail of the presented content shown on the map can be altered. As the vehicle speed increases, the detail shown on the map is reduced, for example, to include only the vehicle's present location in relation to nearby streets and intersections. Some or all of the presentation content, or map directions, may be translated to an audio format and broadcast to the vehicle operator via the speakers 166. In other examples, the size and font of the text, icons, and graphics presented on the display can be altered as the vehicle condition changes, as described with further reference to FIGURE 16. Audio outputs can also be altered, such as by an increase or decrease in audio broadcast based on a change in the vehicle condition. After presenting the output, the processor 168 returns to receiving additional presentation content (block 171). The presentation content can be processed in real-time from the time the content is received over the network by the server to the output of the presentation content to a vehicle operator through the user interface 162. In a further embodiment, the presentation content can be received at one time and stored by the processor 168 before any determination is made of how to present the output. In a still further embodiment, a person other than the vehicle operator can view the presentation content based on a set level of detail.

[0105] FIGURE 16 is a flow chart diagram showing a routine for displaying presentation content based on a vehicle condition. The speed of a vehicle is determined (block 181), for example, by reference to the vehicle speedometer. If the vehicle's speed is less than or equal to 25 miles per hour (mph), the processor 168 generates a detailed display of the presentation content on the display 165 of the user interface 162 (block 182). As the speed of the vehicle increases, the amount of detailed visual information generated for display is reduced. Thus, if the vehicle speed is greater than 25 mph but less than or equal to 40 mph, the processor 168 generates a display of the presentation content with less visual information (block 183) than was generated for the display corresponding to vehicle speeds of less than or equal to 25 mph. When the speed is greater than 40 mph but less than or equal to 70 mph, the processor 168 reduces the amount of detail of the presentation content to be displayed on the display 165, by increasing the font size of the text included in the displayed presentation content and changing the color of the text (block 184). The increased font size and color change provides a display that is easily viewable to the user at high speeds. When the speed is greater than 70 mph, the processor 168 stops generating display information for presentation on the display 165 and switches to a pure audio presentation for output over the speakers 166 (block 185). The specific speed ranges described are used only as an example and should not be construed as limiting. The exact trigger speeds can vary according to various factors, such as road conditions, vehicle specification, or vehicle operator skill, as well as user preferences. Likewise, the output format of the presentation content may be varied according to user interface hardware or software constraints, as well as user preferences.

[0106] In addition to the vehicle example described above, the presentation content can be automatically adjusted based upon one or more vehicle conditions related to a rate of change in vehicle speed, rate of change in vehicle direction, cellular phone or headphone activation, radar detector activation, noise level decibels above a certain value, radio or compact disc usage, television or video disc usage, and turn signals as well as other conditions. The vehicle conditions can be determined by vehicle systems, such as the vehicle speedometer, steering system, engine, braking system, and acceleration system. The vehicle systems can also include additional components, such as a temperature gauge, humidity gauge, wind speed gauge, precipitation gauge, compass, and clock. These additional vehicle system components can be integrated with or separate from the MMS 161, so long as they are capable of being monitored by the MMS 161. Other vehicle conditions, vehicle systems, and components are possible.

[0107] Output of the presentation content can be automatically adjusted based upon one of the vehicle conditions or a combination of the vehicle conditions. For example, the MMS 161 monitors vehicle braking and acceleration systems with reference to an internal clock to determine the rate of change of the vehicle's speed. If the vehicle speed is increasing or decreasing rapidly, the MMS 161 automatically reduces the amount of detail provided in visual presentation content, and can transfer some or all of the presentation content to an audio presentation. In another example, the MMS 161 monitors the movement of the steering wheel and vehicle direction with reference to a compass and a clock to determine a rate of change of the vehicle direction. If the vehicle direction is changing rapidly, the MMS 161 can automatically reduce the amount of detail provided in the visual presentation content, as well as transfer some or all of the presentation content to an audio presentation.

[0108] FIGURE 17 is a block diagram 190 showing a vehicle system for communicating with the system of FIGURE 3A. The vehicle system 191 includes a user interface 192 and a processor 193. The user interface 192 includes a microphone that receives voice instructions from a user and sends the instructions to the processor 193. The processor 193 performs voice processing of the received voice signals and outputs the processed voice signals over a speaker system or a visual display via the user interface 192. The processor 193 can also send the processed voice signals over a wireless network 194 to an operator 196 at an operator system 195.

[0109] FIGURE 18 is a flow chart diagram showing a method for communicating with the system of FIGURE 3A. First, a user inputs voice instructions (block 201) into the vehicle system 191 via the user interface 192. Next, the processor 193 performs voice recognition processing (block 202) on the received voice instructions to determine the content of the instructions. Results from the voice recognition processing is presented to the user (block 203) via the user interface 192 in the form of verbal communication over a speaker system or text messaging on a visual display. The user can determine whether the voice instructions were correctly interpreted (block 204). If the processor 193 has interpreted the inputted voice instructions properly, the user can instruct the vehicle system 191 to send the processed instructions to a predefined destination, such as a server (block 208). Other destinations are possible. Otherwise, the user has an opportunity to contact a human operator to make sure the instructions are properly entered (block 205). If the user decides to contact the human operator, a digest of the interaction between the user and the processor 193 is sent to a human operator 196 (block 206) at an operator system 195. The user is connected to the human operator (block 206), who is tending the operator system 195, including a phone and a computer terminal. The user and the human operator resolve the content of the user-initiated instructions. If the user does not want to be connected to a human operator, the user can return to providing new voice instructions (block 201).

[0110] In a further embodiment, the processor 193 processes a new voice instruction based on a same or similar voice instruction that was previously recorded. For example, on five previous episodes, the user provided a certain voice instruction about a restaurant. The processor 193 made a first determination as to what the user requested for each of the five episodes. In the first four of the five episodes, the user desired directions to the restaurant. In the fifth episode, the user desired stock information about the publicly traded company that owns the restaurant. The approved results, of the processed voice instruction, including four requests for directions and one request for stock information are recorded by the processor 193 with the voice instruction for each episode. The next time the user provides that same voice instruction, the processor 193 can use the previously recorded approved results for processing the new voice instruction.

[0111] Various methods of using the previously recorded approved results can be performed. In a first method, the most frequently recorded result becomes the result of the newly provided specific voice instruction. So, using the example above, directions are provided to the user, because the recorded results show requests for directions to the restaurant occurring more frequently than requests for stock information, or other types of requests. In a second method, a most recent recorded result becomes the result of the newly provided specific voice instruction. So, using the example above, stock information is provided to the user, because the last recorded result was a request for stock information.

[0112] With respect to the first four episodes described above, the processor 193 processes the voice instruction as either a request for directions or as a request for something other than directions. If the voice instruction was processed as a request for directions, the user can approve of the processing, and the resulting request for directions to the restaurant is recorded with the voice instruction for that episode. If the voice instruction was processed as a request for something other than directions, the user can make a further request to the processor 193 to provide directions, thereby indicating that the processing of the voice instructions was incorrect.

[0113] The user can indicate that the processing was incorrect in a number of different ways. A user can correct the processing by contacting a human operator (block 205). The human operator can make the change for that episode and the change is recorded. A correction can also be made by submitting a subsequent voice instruction request to the processor 193 for reinterpreting the original voice instruction. Returning to the previous example, once the voice instruction has been corrected by changing to a request for directions and the user approves of the change, the result of the directions to the restaurant is recorded with the voice instruction for that episode.

[0114] Communication between a driver of a vehicle and an agent at a call center can occur through messaging. As the agent and the driver communicate, their voices are captured, digitized, compressed and sent to a server. The server routes the voice segments to a call center agent desktop application, which includes a human call center agent sitting in front of an agent console, including a screen, keyboard, speaker, microphone, and headset. The agent can interact with multiple vehicles, each vehicle represented by a session on the agent's screen. Each communication session is represented by a window, that displays icons representing the received utterances from the vehicle.

[0115] The agent can listen to a specific utterance by selecting the corresponding icon with a mouse. In a further embodiment, the voice from the driver of the vehicle is first processed by a speech to text recognition engine then placed as text into that window.

[0116] To reply to a specific driver of a vehicle, the agent types in the corresponding window for that vehicle. The typed text is converted to voice output by a text to speech server, compressed, and sent back to the vehicle. The typed text is first stored in the TCU, then decompressed and played to the driver. In a further embodiment, the operator can speak the reply rather than typing a response. Both the driver and the operator can listen to an utterance of the message again by asking their respective device to replay the utterance.

[0117] The agent can interact with multiple vehicles simultaneously. A log of each conversational session can be kept by the agent at the call center, as well as by the driver of the vehicle participating in that conversation. The agent or the driver can listen to previously stored instructions. The agent can be much more efficient since he does not need to ask

the driver to repeat the utterance or to apologize for putting the driver on hold (the efficiency is implied by the half duplex message oriented nature of the communication.)

[0118] Agent messaging allows agents to multitask between several customers simultaneously using multiple windows. The customers can each communicate with an agent through voice media. The agents receive the voice messages from multiple customers, which are queued up for the agents to peruse in sequence. Other related features are possible, such as limited agent control of other client applications.

[0119] Communication between a customer and a call center agent can include two application programs: one for the customer (TCU), client, and one for the agent, server. The server is designed for store-and-forward message passing between individual customers and agents. The agent's application program can include some form of Windows, executable with a non-browser GUI. The customer application program is an application running on the TCU. Messages from the customers to the agents can be audio recordings, while messages from the agents to the customers are primarily text messages passed through a text-to-speech engine. However, other methods of communication are possible, including allowing agents to send audio messages and customers to send text messages.

[0120] For example, a customer contacts an agent because he is lost and wants directions to a specific location. The customer's position and recent movements can be uploaded for review by the agent. Based on the customer's position, the agent can control a navigation application and provide directions to the specific location.

[0121] In a further example, a customer's vehicle has failed or was involved in an accident; however, emergency services do not need to be summoned. The vehicle information is uploaded based on a connection to a CAN bus. The customer position is also uploaded.

[0122] In yet a further example, an agent can perform a concierge service for the customer, which can be handled quickly, such as securing tickets to a show. Obtaining the tickets can often take under five minutes; however, other times are possible. External services, such as a ticket Web page can be used by the agent via the agent console. The agent can access the customer's Personal Digital Assistant (PDA) or other form of electronic scheduling device through synchronization to determine the customer's availability to attend the show. The PDA can include information such as calendar items, addresses, or phone numbers. The agent can also access the customer's preferences and information, such as credit card information and email for conducting the service. Once the tickets are obtained, a confirmation of the service is sent to the customer via an email.

[0123] In yet a further example, an agent can perform a concierge service for the customer, which takes a long time, such as placing reservations at a restaurant that is not answering their phone. Services that take over 30 minutes to complete can be considered long-term services. Other time limits can be used. The agent follows the same steps for the short-term concierge services; however, a pause session is usually required due to the length of time required to complete the service. The agent can initiate a session restart once the service has been completed.

[0124] Communication between the customer and the agent can include transmitting a new customer voice request, position, and vehicle information to the agent. The agent can have control of other customer vehicle applications at the same time as the receipt of the new customer's voice request. During each session, the agent can use external services via the agent console. Access to each customer's PDA can be obtained through synchronization. The agents can view customer data from the PDA, including phone numbers, schedule items, to-do items, and other data. Customer preferences can also be obtained to assist in completing the request. Once the request is complete, the agent can send a confirmation, such as a summary notice by email. If the request requires a large period of time to complete, client disconnection can occur with intermittent reconnection to check for delayed responses.

[0125] Message requests from multiple clients can be queued for handling by a single agent. A client for the customer and a client for the agent can communicate with a single server. The agent sends back a reply message, which is queued for later playback, when the driver is ready. Problematic voice-recognition situations to the agent, such as unrecognizable voice data, can be automatically handed-off to an agent who is informed of the context. Data can be considered unrecognizable if the data is not understood correctly after three attempts. Other classifications of unrecognizable data are possible.

[0126] The agent can access and control each customer profile. Incoming requests from customers can be prioritized by factors, such as time constraints, as well as other factors. For example, the agent can designate customer X as needing a quick turnaround of all responses. Customer X's requests are inserted close to the head of the queue rather than at the end.

[0127] The agent application program can include scripts, shortcuts, and other tools for providing common responses, such as "I'm sorry sir, there are no tickets for that performance. Is there another time which might also work?" The agent can annotate the communication for review by themselves or others, or to refresh their context when restarting a paused session.

[0128] The communication between the customer and the agent can be archived for some period of time. The archived communication can be used to review or make a copy of old communication. If available, the agent can invisibly serve as a voice recognition tool for another application. For example, a customer requests weather information for a city. After a second failed attempt to recognize the data, the agent receives the audio for the customers' third attempt with

the annotation for a city.

[0129] If the customer is arranging a scheduled item, such as reservations for the theatre or another event, the agent can automatically check for conflicts with the customer's schedule on the customer's PDA. The agent may not know what the conflicting item is, only that a conflict exists.

5 **[0130]** An agent starts the agent application program, using an agent identifier, such as a login name or other identifier, recognizable by the server. The agent application program manages a queue of lines of communication with multiple customers. When a customer wishes to talk with an agent, the client application negotiates with the server to establish a session with a particular agent. Each customer may have a regular or default agent who is familiar with that customer, though the workload should be balanced across all working agents. The agent's work includes pushing a 'next message' button, which brings up the next message in the queue. When pressed, the agent gets the most recent audio from the customer presented to them. On the agent's screen is a transcript of the conversation with this customer. The transcript is a combination of the customer's audio recordings and the agent's text-based responses and commands. The transcript of the communication will likely be interleaved, so the application program will allow the agent to select specific lines of the transcript. If the transcript is a text message, the full text will be displayed, or if the transcript is audio communication, 10 the customer's audio recoding will be replayed. The transcript can be scrolled back indefinitely, even to a previous session on a previous day, week, or as storage permits, as well as other times. Besides the transcript, other information about the queue is available, such as how many other customers are waiting, average wait time, or relative priorities of the customers in the queue, as well as other information.

15 **[0131]** The agent can also have information about the customer from a database available through the agent's console. The information can include, for example, customer and family names, preferred salutations, and addresses of important locations, as well as other information. The agent application program allows agents to respond to customers, either to enter text directly, to respond with shortcuts that have common phrases, to incorporate the customer's profile information, and to respond with commands that may cause functional changes on the client machine, such as by specifying a new destination for the navigation component or by downloading an address to their PDA, as well as other types of responses. 20 The agent can modify the customer's relative priority in the queue. Thus, based on how quickly the customer needs a response the message can enter the queue at a higher priority and thus skip ahead of other customers. For example, customer A has car engine difficulties while customer B wants to purchase a compact disc. The request from customer A enters the agent's queue at a higher priority than the request from customer B.

25 **[0132]** While the invention has been particularly shown and described as referenced to the embodiments thereof, those skilled in the art will understand that the foregoing and other changes in form and detail may be made therein without departing from the scope of the invention.

APPENDIX

5

To specify an alert message *string*, use:

ALERT *string*

10

To specify an unconditional branch to *label*, use:

GOTO *label*

15

To specify a conditional *cond* branch to *label*, use:

IF *cond* THEN GOTO *label*

20

To specify a label *string* for a branch, use:

LABEL *string*

25

To terminate execution, use:

QUIT

30

To specify a synthesized speech utterance *string*, use:

SAY *string*

To specify a script name *string*, use:

35

SCRIPTNAME *string*

To specify a variable name *varname*, use:

40

SET *varname*

To specify a subscript *scriptname*, use:

45

SUBSCRIPT *scriptname*

To specify a wait condition *string*, use:

50

WAITFOR *string* <YES|NO|DATE|PHONENUM|CUSTID>

Claims

55

1. A system (30) for providing a multi-modal communications infrastructure for automated call center (11) operation, comprising:

a telephony interface (35) configured to accept a multi-modal call from a caller (12-15);

a messaging server (31) configured to automatically assign the multi-modal call through a session manager (47) to a session under operation of a in the multi-modal call live agent (81) and to receive voice communications and text messages that simultaneously originate from the caller;
 a speech recognition module (36) configured to convert incoming speech (54) from the voice communications in the multi-modal call into transcribed text (57);
 a text message processor configured to match at least one of the text messages from the caller with the transcribed text (57) of the incoming speech (54) during the session that is on-going; and
 an agent console (39) configured to progressively process the transcribed text (57) and the incoming text messaging during the session through an agent application (43) by performing a customer support scenario (22) interactively monitored and controlled by the live agent (81).

2. A system (30) according to Claim 1, further comprising at least one of:

a database (34) configured to store the incoming text messaging and the transcribed text (57) as a log entry (44); and
 a post-processing module configured to perform post-processing on the incoming text messaging and the transcribed text (57).

3. A system (30) according to Claim 1, further comprising:

a caller identification module configured to assign a caller identification (71) to the incoming text messaging; and
 the messaging server (31) configured to forward the incoming text messaging to the live agent (81) based on the caller identification (71).

4. A system (30) according to Claim 1, further comprising:

an identification (71) module configured to flag the incoming text messaging as verbatim caller data; and
 a database (34) configured to store the flagged incoming text messaging.

5. A method for providing a multi-modal communications infrastructure for automated call center (11) operation, comprising:

accepting a multi-modal call from a caller (12-15) through a telephony interface (35) ; ;
 automatically assigning the multi-modal call through a session manager (47) to a session under operation of a live agent (81);
 in the multi-modal call receiving voice communications and text messages in the multi-modal call that simultaneously originate from the caller;
 converting incoming speech (54) from the voice communications in the multi-modal call into transcribed text (57);
 matching at least one of the text messages from the caller with the transcribed text (57) of the incoming speech (54) during the session that is on-going; and
 progressively processing the transcribed text (57) and the incoming text messaging during the call session through an agent application (43) by performing a customer support scenario (22) interactively monitored and controlled by the live agent (81).

6. A method according to Claim 5, further comprising at least one of

storing the incoming text messaging and the transcribed text (57) as a log entry (44); and
 performing post-processing on the incoming text messaging and the transcribed text (57).

7. A method according to Claim 5, further comprising:

assigning a caller identification (71) to the incoming text messaging; and
 forwarding the incoming text messaging to the live agent (81) based on the caller identification (71).

8. A method according to Claim 5, further comprising:

flagging the incoming text messaging as verbatim caller data; and
 storing the flagged incoming text messaging.

Patentansprüche

1. Ein System (30) zum Bereitstellen einer multi-modalen Kommunikationsinfrastruktur für automatisierten Callcenter (11)-Betrieb, aufweisend:

ein Telefonie-Interface (35), das konfiguriert ist, einen multi-modalen Anruf von einem Anrufer (12-15) zu empfangen;
 einen Messaging-Server (31), der konfiguriert ist, automatisch den multi-modalen Anruf durch einen Session-Manager (47) zu einer Session unter dem Betrieb eines Live-Agenten (81) zuzuweisen und Sprachkommunikation und Textnachrichten in dem multi-modalen Anruf zu empfangen, die simultan von dem Anrufer stammen;
 ein Spracherkennungsmodul (36), das konfiguriert ist, eingehende Sprache (54) von der Sprachkommunikation in dem multi-modalen Anruf in transskribierten Text (57) zu konvertieren;
 einen Textnachricht-Prozessor, der konfiguriert ist, zumindest eine der Textnachrichten von dem Anrufer mit dem transskribierten Text (57) der eingehenden Sprache (54) während der laufenden Session zu matchen; und
 eine Agenten-Konsole (39), die konfiguriert ist, progressiv den transskribierten Text (57) und das eingehende Textmessaging während der Session durch eine Agentenapplikation (43) zu verarbeiten, indem ein Kunden-Support-Szenario (22) ausgeführt wird, das interaktiv überwacht und durch den Live-Agenten (81) gesteuert wird.

2. Ein System (30) nach Anspruch 1, ferner aufweisend zumindest eines von:

einer Datenbank (34), die konfiguriert ist, das eingehende Textmessaging und den transskribierten Text (57) als Log-Eintrag (44) zu speichern; und
 ein Post-Processing-Modul, das konfiguriert ist, ein Post-Processing auf dem eingehenden Textmessaging und dem transskribierten Text (57) durchzuführen.

3. Ein System (30) nach Anspruch 1, ferner aufweisend:

ein Anrufer-Identifikationsmodul, das konfiguriert ist, eine Anrufer-Identifikation (71) dem eingehenden Textmessaging zuzuweisen; und
 der Messaging-Server (31) ist konfiguriert, das eingehende Textmessaging an den Live-Agenten (81) basierend auf der Anruferidentifikation (71) weiterzuleiten.

4. Ein System (30) nach Anspruch 1, ferner aufweisend:

ein Identifikations (71)-Modul, das konfiguriert ist, das eingehende Textmessaging als wörtliche Anruferdaten zu flaggen; und
 eine Datenbank (34), die konfiguriert ist, das geflaggte eingehende Textmessaging zu speichern.

5. Ein Verfahren zum Bereitstellen einer multi-modalen Kommunikationsinfrastruktur für automatisierten Callcenter (11)-Betrieb, aufweisend:

Annehmen eines multi-modalen Anrufs von einem Anrufer (12-15) durch ein Telefonie-Interface (35);
 automatisches Zuweisen des multi-modalen Anrufs durch einen Session-Manager (47) zu einer Session unter Betrieb eines Live-Agenten (81);
 Empfangen von Sprachkommunikation und Textnachrichten in dem multi-modalen Anruf, die simultan von dem Anrufer stammen;
 Konvertieren eingehender Sprache (54) von der Sprachkommunikation in dem multi-modalen Anruf in transskribierten Text (57);
 Matchen zumindest einer der Textnachrichten von dem Anrufer mit dem transskribierten Text (57) der eingehenden Sprache (54) während der Session, die andauert; und
 Progressives Verarbeiten des transskribierten Textes (57) und des eingehenden Textmessaging während der Anrufer-Session durch eine Agentenapplikation (43) durch Durchführen eines Kunden-Support-Szenarios (22), das interaktiv überwacht und durch den Live-Agenten (81) gesteuert wird.

6. Ein Verfahren nach Anspruch 5, ferner aufweisend zumindest eines von:

Speichern des eingehenden Textmessaging und des transskribierten Textes (57) als Log-Eintrag (44); und
 Durchführen von Post-Processing an den eingehenden Textmessaging und dem transskribierten Text (57).

7. Ein Verfahren nach Anspruch 5, ferner aufweisend:

Zuweisen einer Anruferidentifikation (71) zu dem eingehenden Textmessaging; und Weiterleiten des eingehenden Textmessaging an den Live-Agenten (81) basierend auf der Anruferidentifikation (71).

8. Ein Verfahren nach Anspruch 5, ferner aufweisend:

Flaggen des eingehenden Textmessaging als wörtliche Anruferdaten; und Speichern des geflaggtten eingehenden Textmessaging.

Revendications

1. Système (30) destiné à fournir une infrastructure de communications multimodales pour le fonctionnement automatisé d'un centre d'appels (11) comprenant :

une interface de téléphonie (35) configurée pour accepter un appel multimodal en provenance d'un appelant (12-15) ;

un serveur de messagerie (31) configuré pour attribuer de manière automatique l'appel multimodal par l'intermédiaire d'un gestionnaire de session (47) à une session sous la gestion d'un interlocuteur direct (81) et pour recevoir des communications vocales et des messages textuels dans l'appel multimodal qui proviennent simultanément de l'appelant ;

un module de reconnaissance vocale (36) configuré pour convertir une voix entrante (54) à partir des communications vocales dans l'appel multimodal en un texte transcrit (57) ;

un processeur de message textuel configuré pour mettre en correspondance l'un au moins des messages textuels en provenance de l'appelant avec le texte transcrit (57) de la voix entrante (54) pendant la session en cours ; et

une console d'interlocuteur (39) configurée pour traiter de manière progressive le texte transcrit (57) et la messagerie textuelle entrante au cours de la session par l'intermédiaire d'une application d'interlocuteur (43) en exécutant un scénario de support client (22) surveillée de manière interactive et commandée par l'interlocuteur direct (81).

2. Système (30) selon la revendication 1, comprenant en outre l'un au moins :

d'une base de données (34) configurée pour stocker la messagerie textuelle entrante et le texte transcrit (57) sous la forme d'une entrée de journal (44) ; et

d'un module de post-traitement configuré pour exécuter un post-traitement sur la messagerie textuelle entrante et le texte transcrit (57).

3. Système (30) selon la revendication 1, comprenant en outre :

un module d'identification d'appelant configuré pour attribuer une identification d'appelant (71) à la messagerie textuelle entrante ; et

le serveur de messagerie (31) étant configuré pour transférer la messagerie textuelle entrante à l'interlocuteur direct (81) sur la base de l'identification de l'appelant (71).

4. Système (30) selon la revendication 1, comprenant en outre :

un module d'identification (71) configuré pour signaler la messagerie textuelle entrante en tant que données d'appelant mot à mot ; et

une base de données (34) configurée pour stocker la messagerie textuelle entrante signalée.

5. Procédé destiné à fournir une infrastructure de communications multimodales pour le fonctionnement automatisé d'un centre d'appels (11) comprenant :

l'acceptation d'un appel multimodal en provenance d'un appelant (12-15) par l'intermédiaire d'une interface de téléphonie (35) ;

l'attribution automatique de l'appel multimodal par l'intermédiaire d'un gestionnaire de session (47) à une session sous la gestion d'un interlocuteur direct (81) ;
la réception des communications vocales et des messages textuels dans l'appel multimodal qui proviennent simultanément de l'appelant ;
5 la conversion d'une voix entrante (54) à partir des communications vocales dans l'appel multimodal en un texte transcrit (57) ;
la mise en correspondance de l'un au moins des messages textuels en provenance de l'appelant avec le texte transcrit (57) de la voix entrante (54) pendant la session en cours ; et
10 le traitement progressif du texte transcrit (57) et de la messagerie textuelle entrante au cours de la session d'appel par l'intermédiaire d'une application d'interlocuteur (43) en exécutant un scénario de support client (22) surveillée de manière interactive et commandée par l'interlocuteur direct (81).

6. Procédé (30) selon la revendication 5, comprenant en outre l'un au moins :

15 du stockage de la messagerie textuelle entrante et du texte transcrit (57) sous la forme d'une entrée de journal (44) ; et
de l'exécution d'un post-traitement sur la messagerie textuelle entrante et le texte transcrit (57).

7. Procédé selon la revendication 5, comprenant en outre :

20 l'attribution d'une identification d'appelant (71) à la messagerie textuelle entrante ; et
le transfert de la messagerie textuelle entrante à l'interlocuteur direct (81) sur la base de l'identification de l'appelant (71).

8. Procédé selon la revendication 5, comprenant en outre :

25 le signalement de la messagerie textuelle entrante en tant que données d'appelant mot à mot ; et
le stockage de la messagerie textuelle entrante signalée.

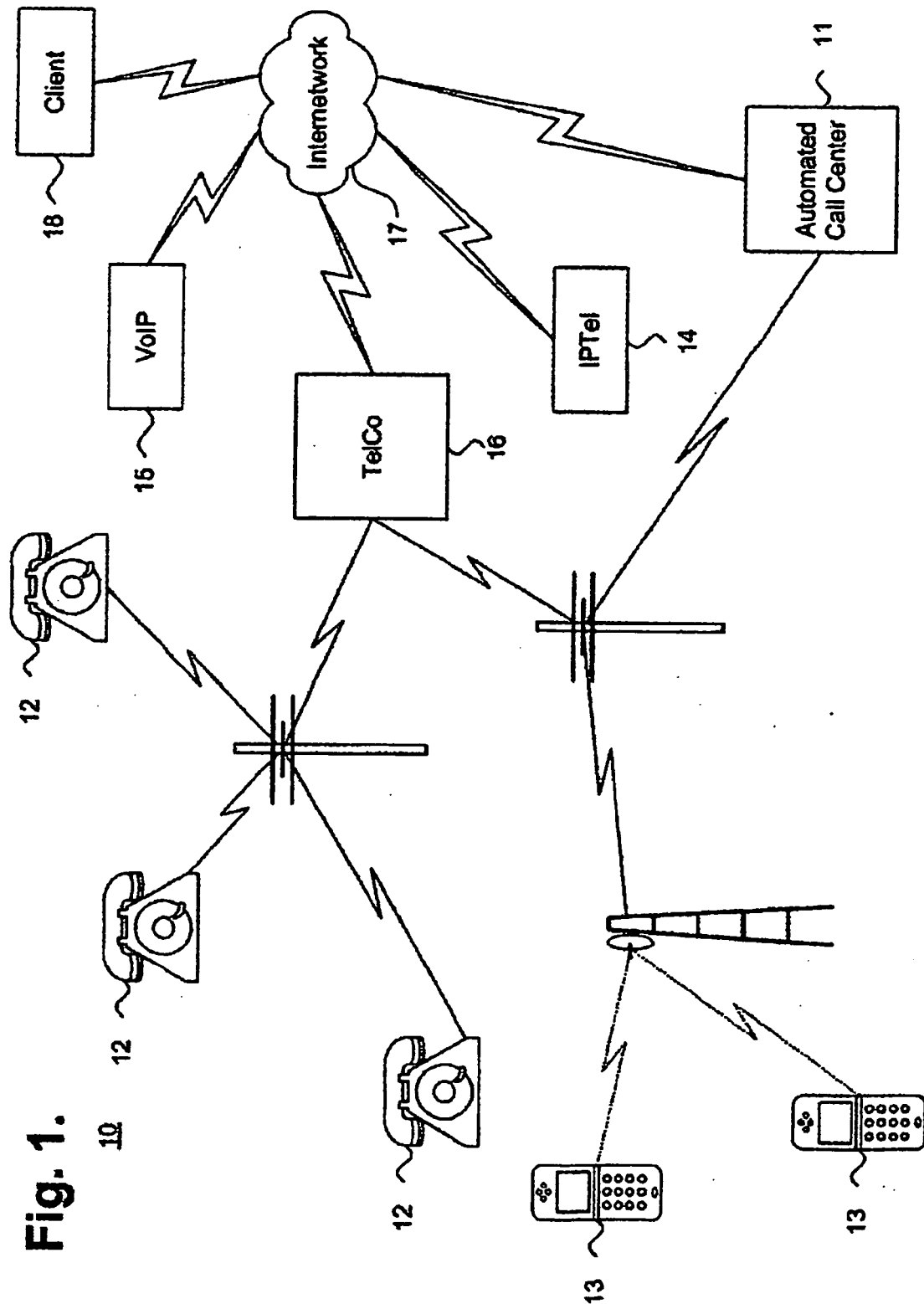


Fig. 1.

Fig. 2.

20

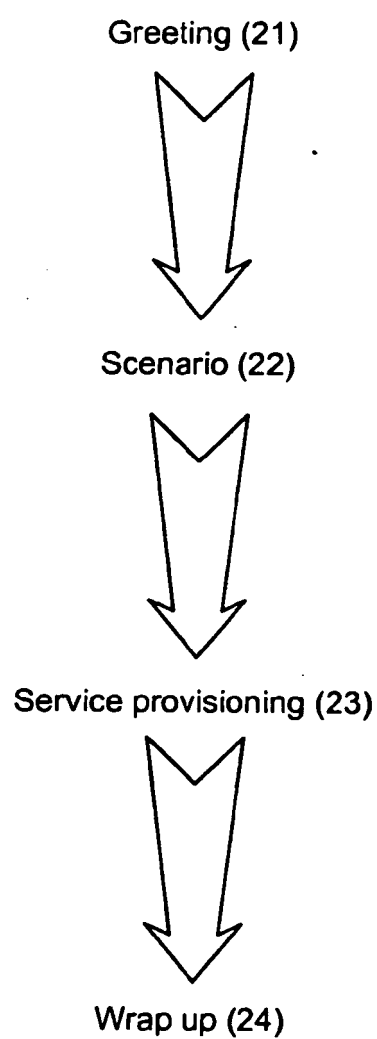


Fig. 3A.

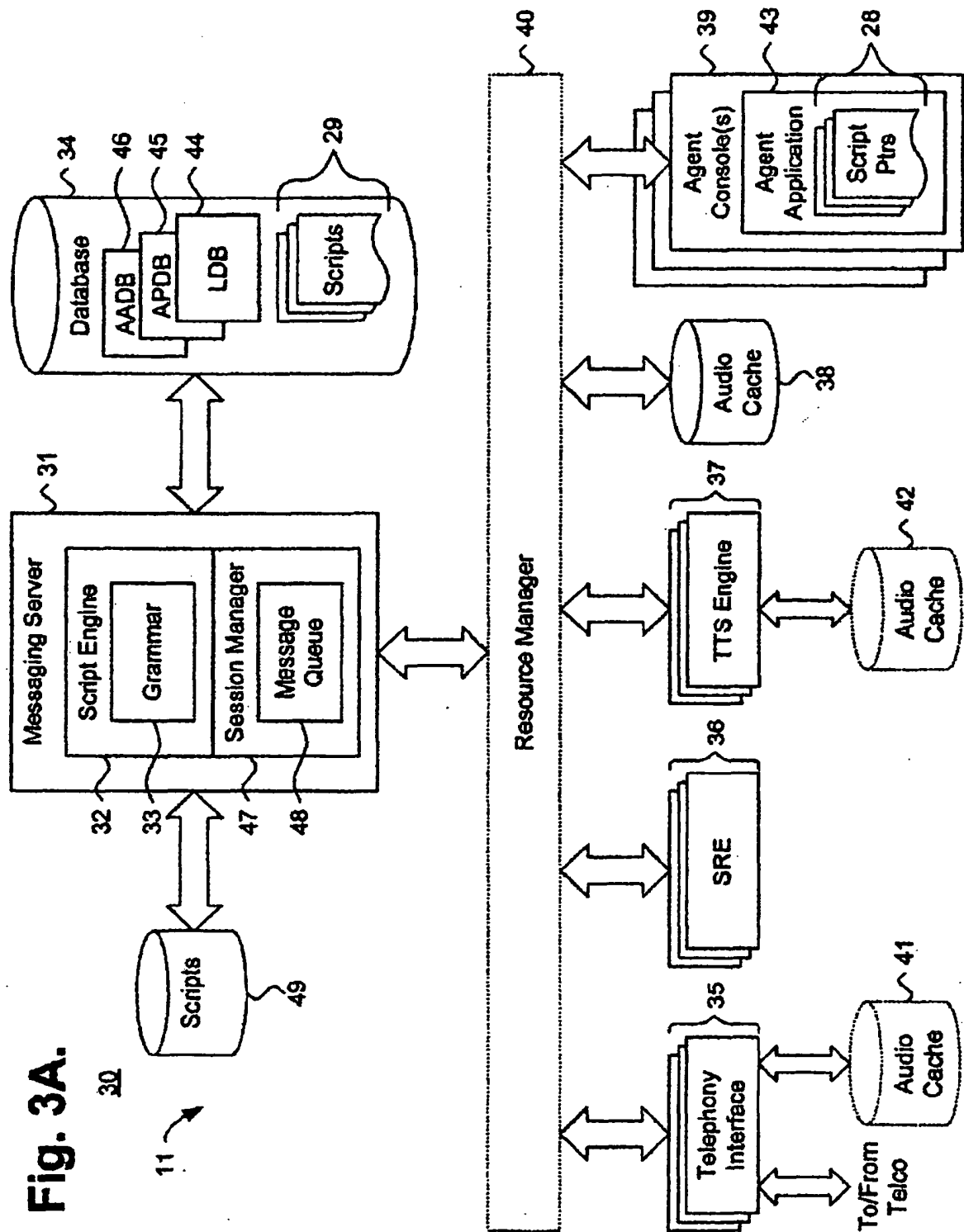


Fig. 3B.

11
160

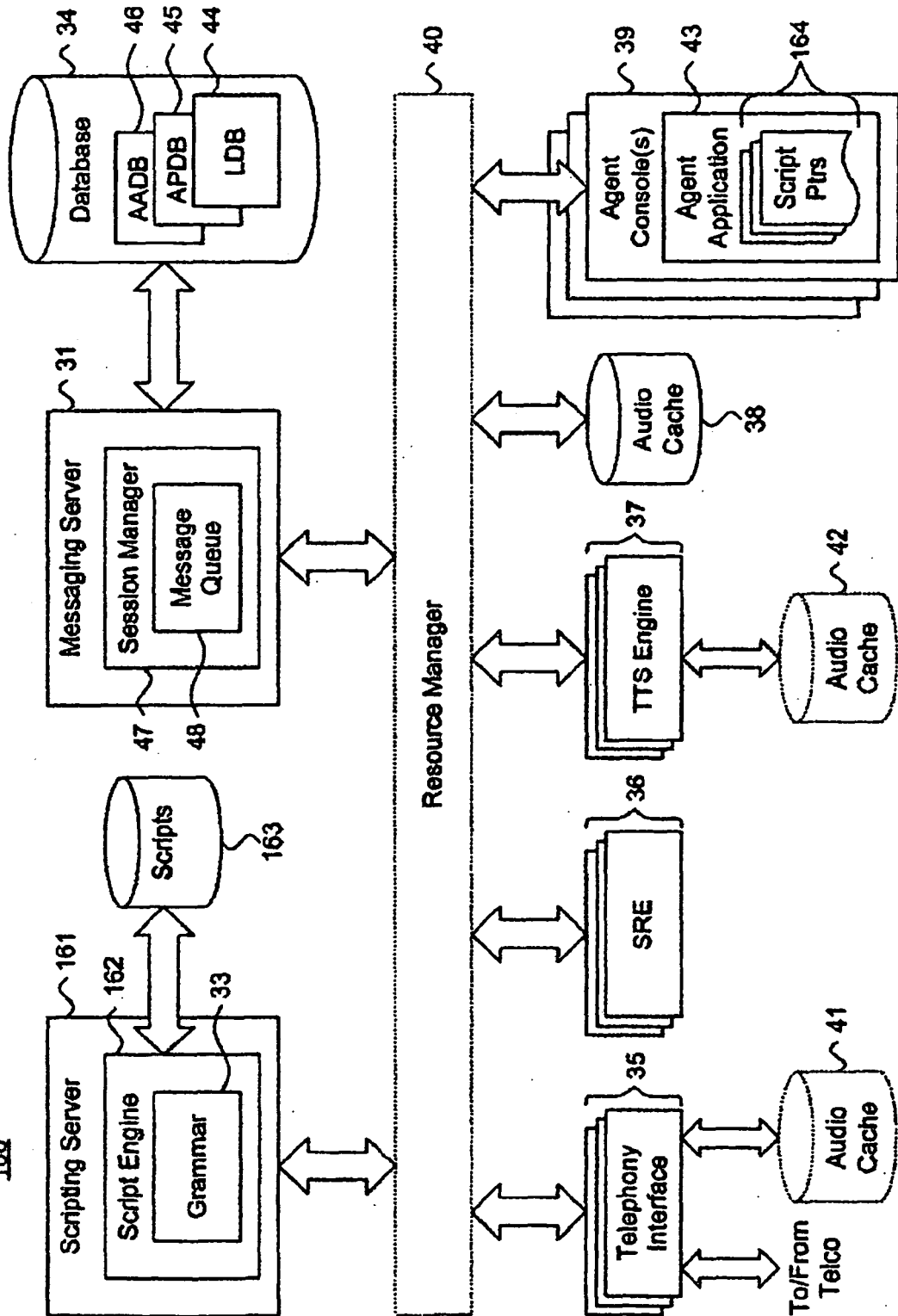


Fig. 4.

50

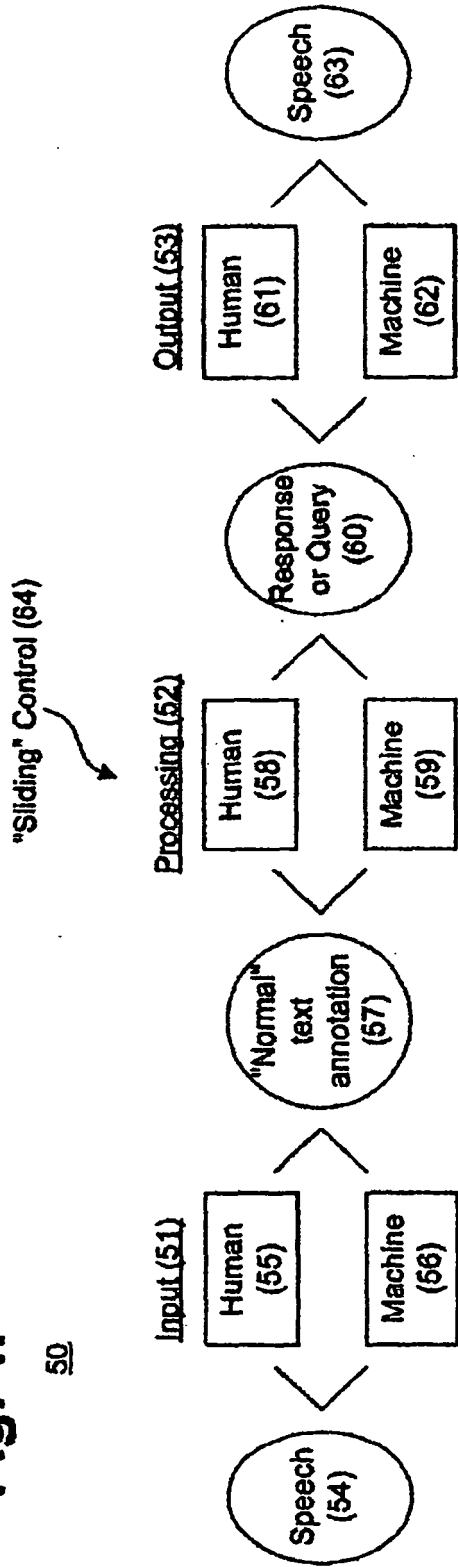


Fig. 5.

70

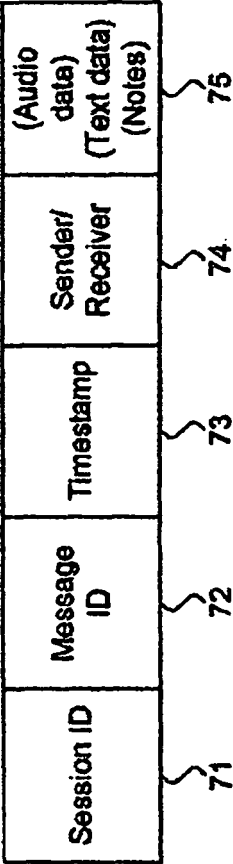


Fig. 6.

80

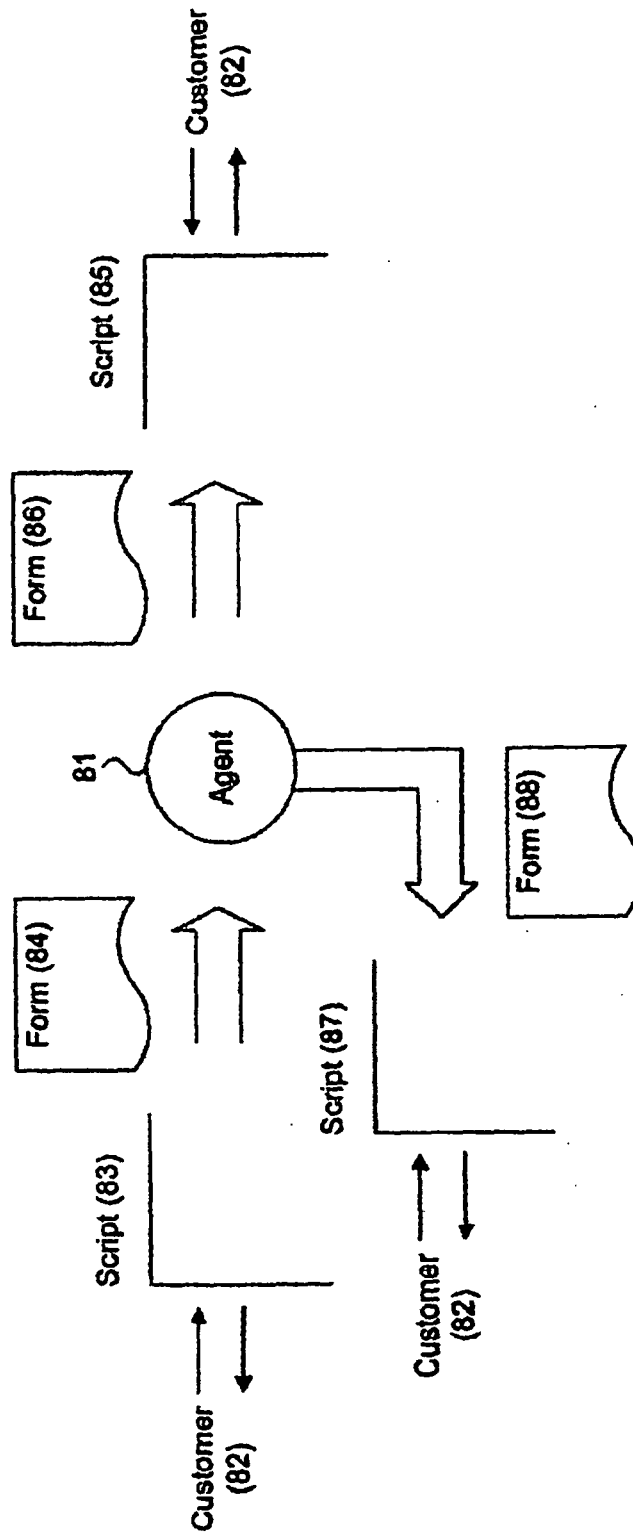


Fig. 7A.

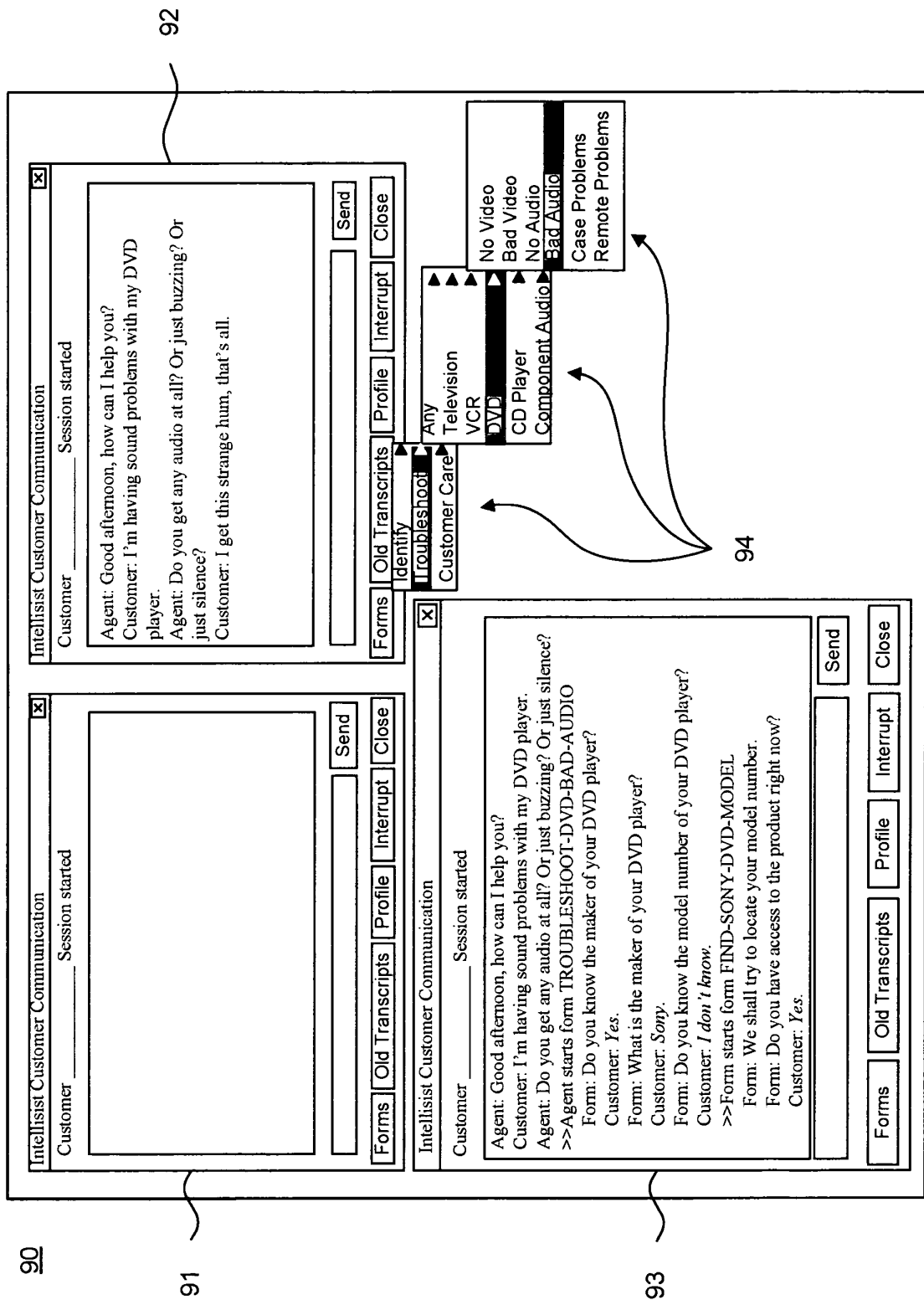


Fig. 7B.

95

Guide Console – Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Search

Address

96

Caller 1	Caller 2	Caller 3	Caller 4	Caller 5	Caller 6
CALL TIME : Transfer to Agent End Call					
Caller Phone Listen Update			Reason for Call Listen Ask Again Update		
Phone Number			Select from the list below.		
Key Code Listen Update			☐ Order (Ctrl+1)		
Yellow Box			☐ Order Status (Ctrl+2)		
Customer Number Listen Update			☐ Credit Card Status (Ctrl+3)		
Pink Box			☐ Billing Question (Ctrl+4)		
			☐ Item Availability (Ctrl+5)		
			☐ Request Catalog (Ctrl+6)		
			☐ Customer Service/Want Live Agent/Other (Ctrl+7)		
Caller Profile Listen Update			Address		
Zip Code			Address 1		
Last Name			Address 2		
First Name			City		

97

98

99

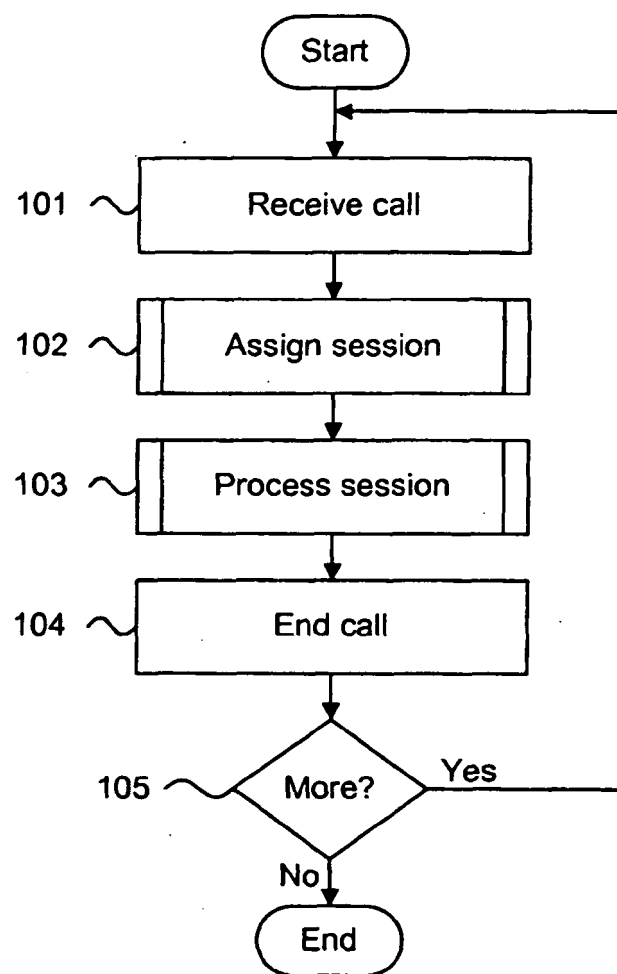
Fig. 8.100

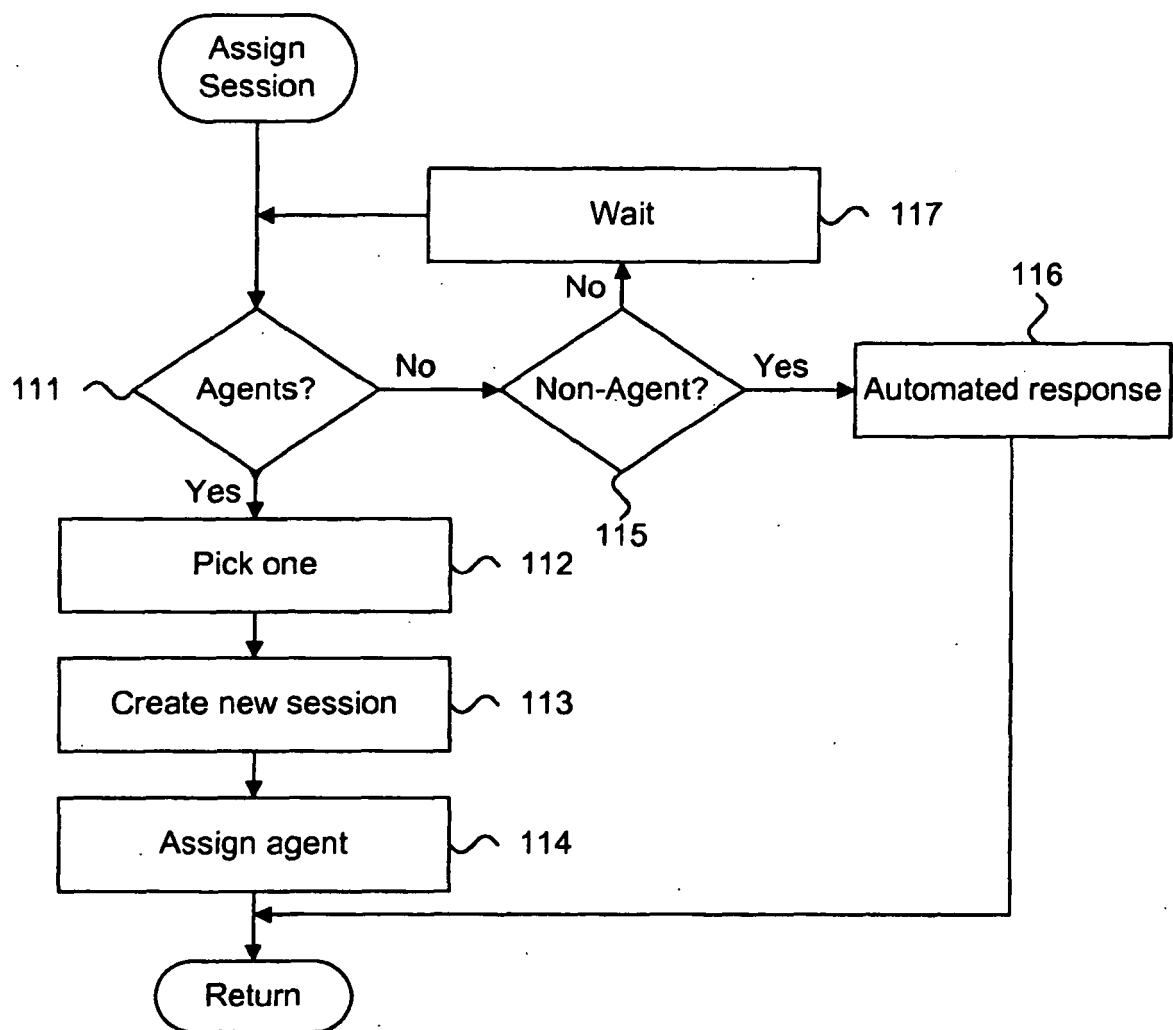
Fig. 9.110

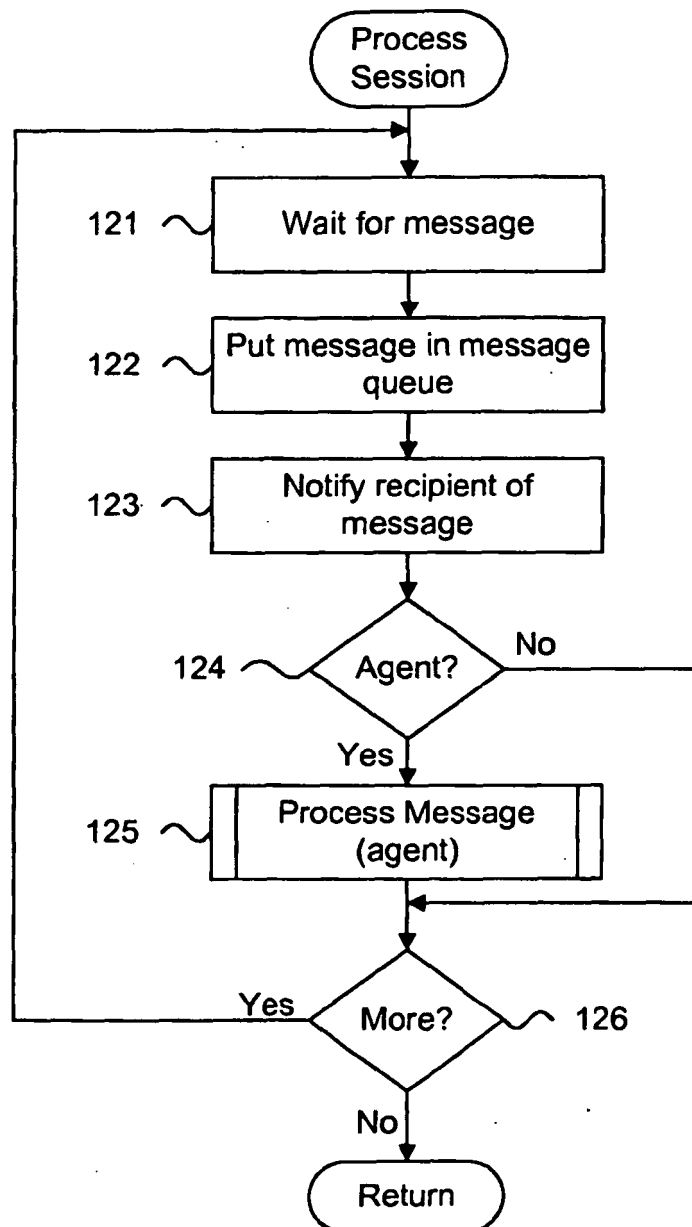
Fig. 10.120

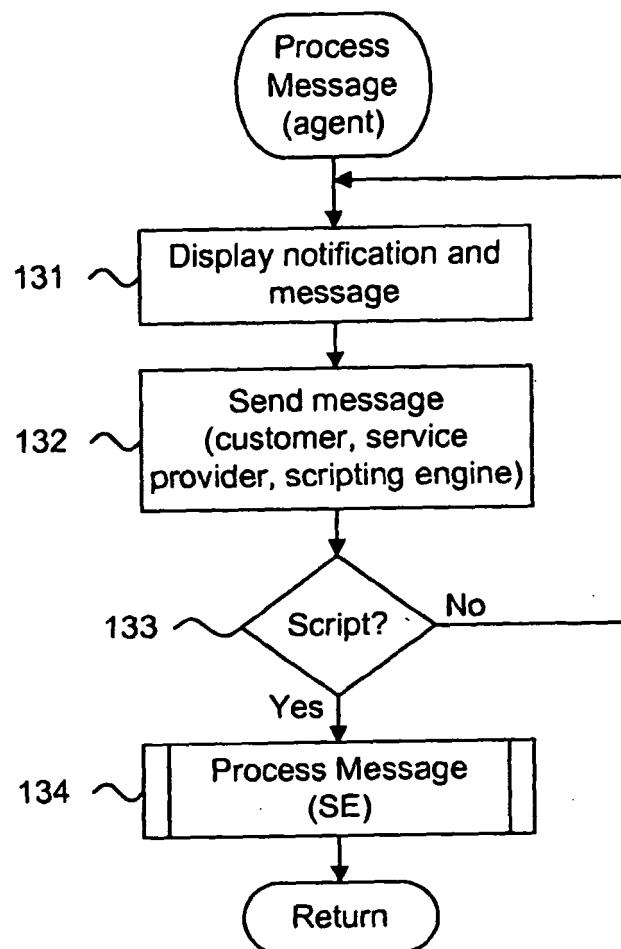
Fig. 11.130

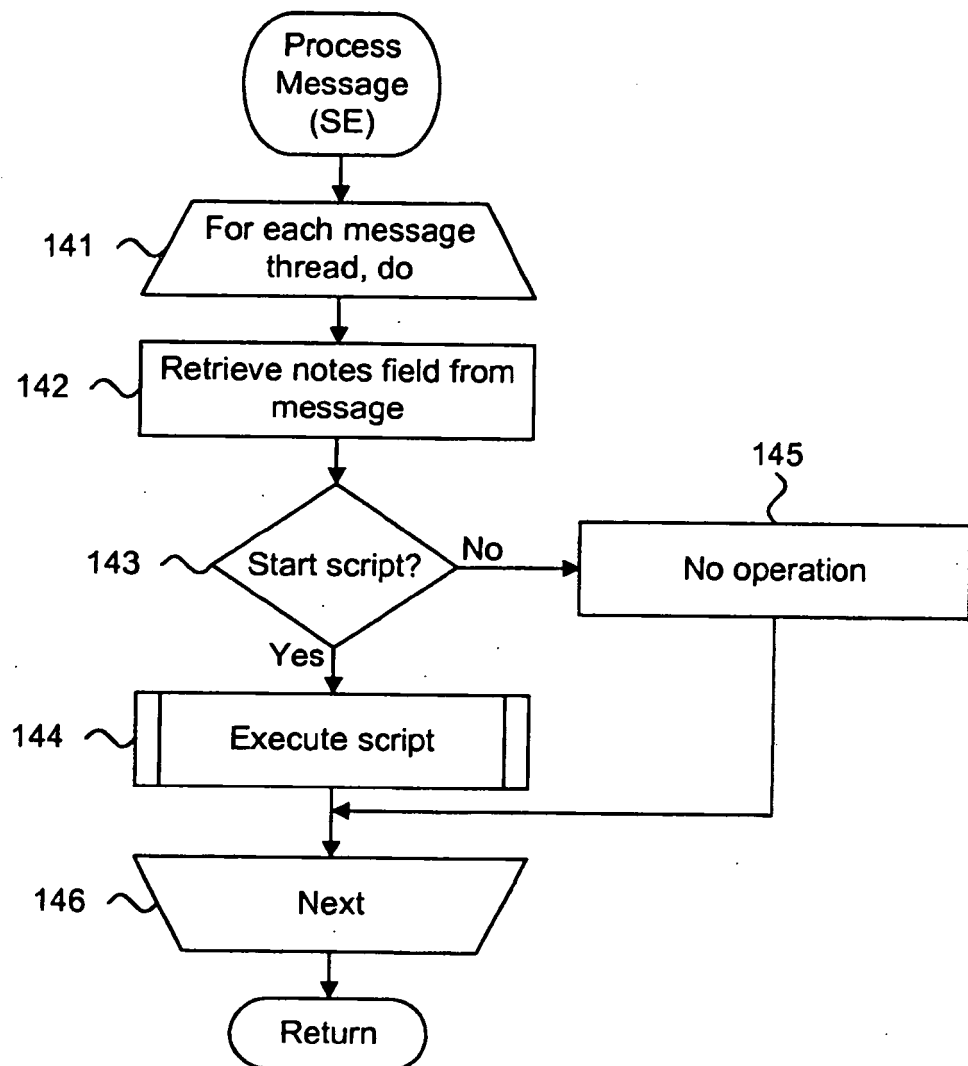
Fig. 12.140

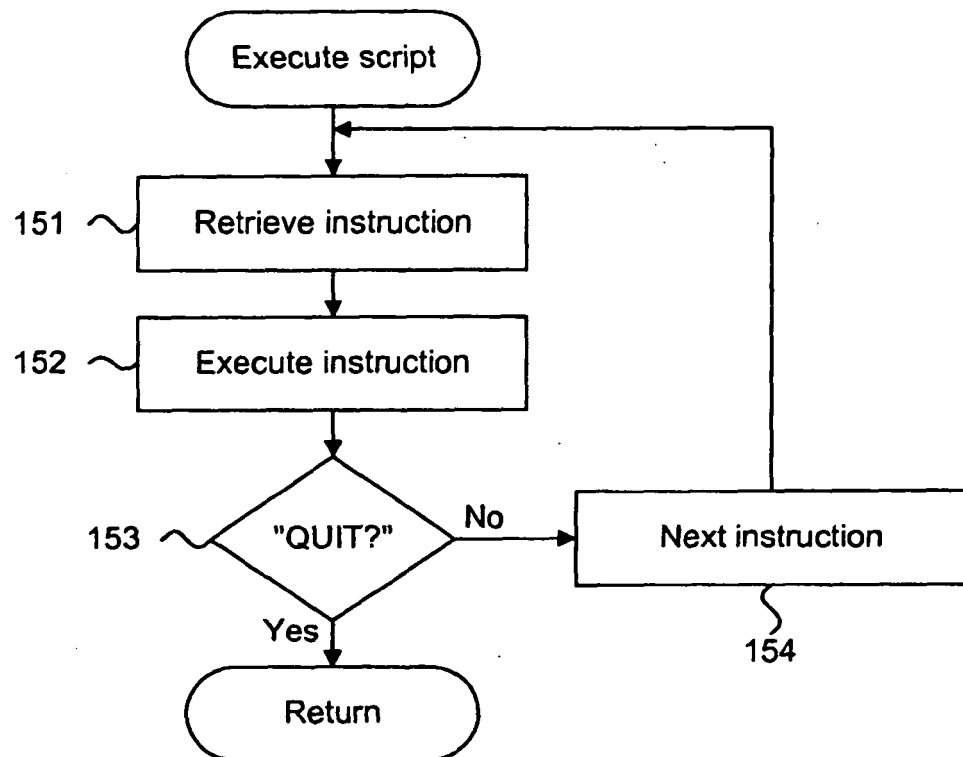
Fig. 13.150

Fig. 14.

160

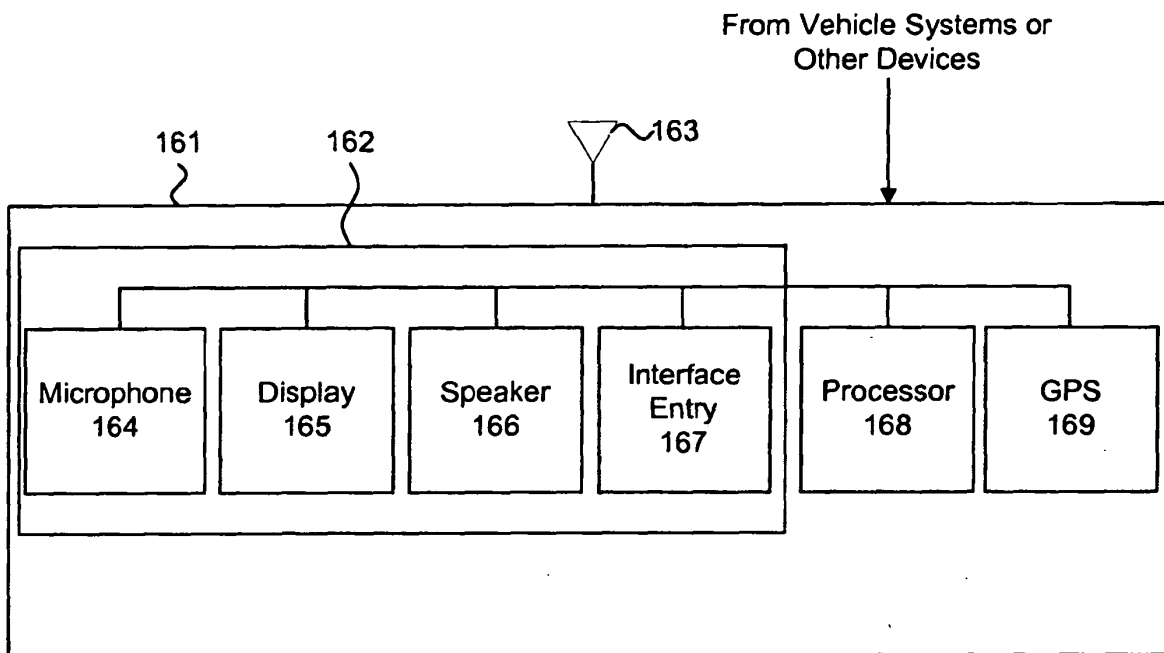


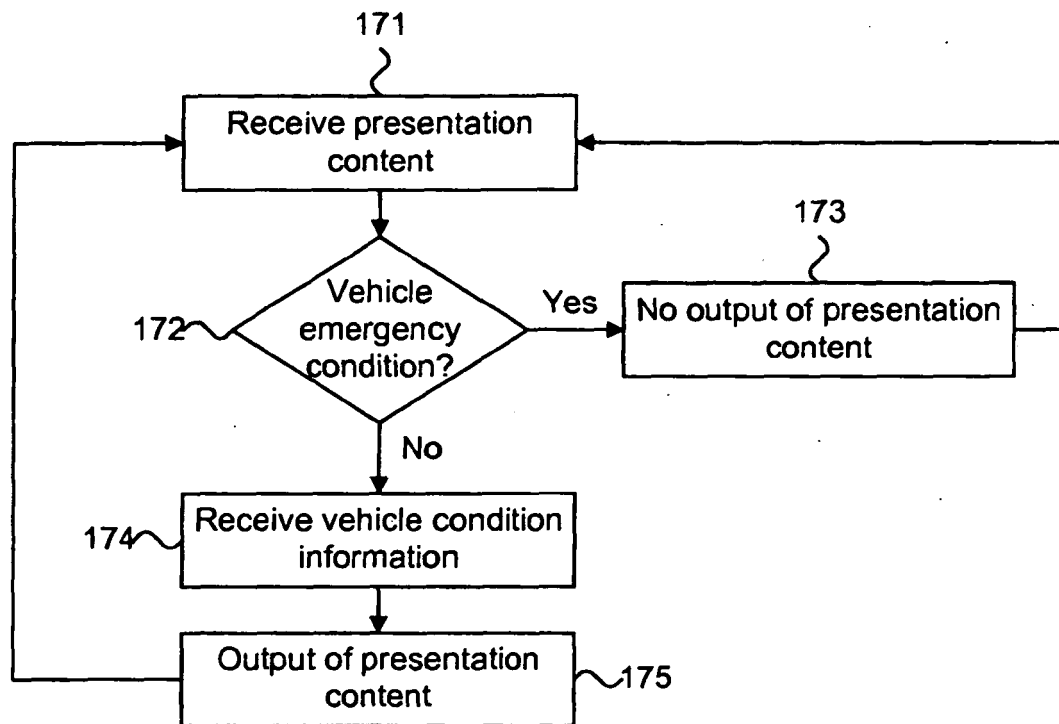
Fig. 15.170

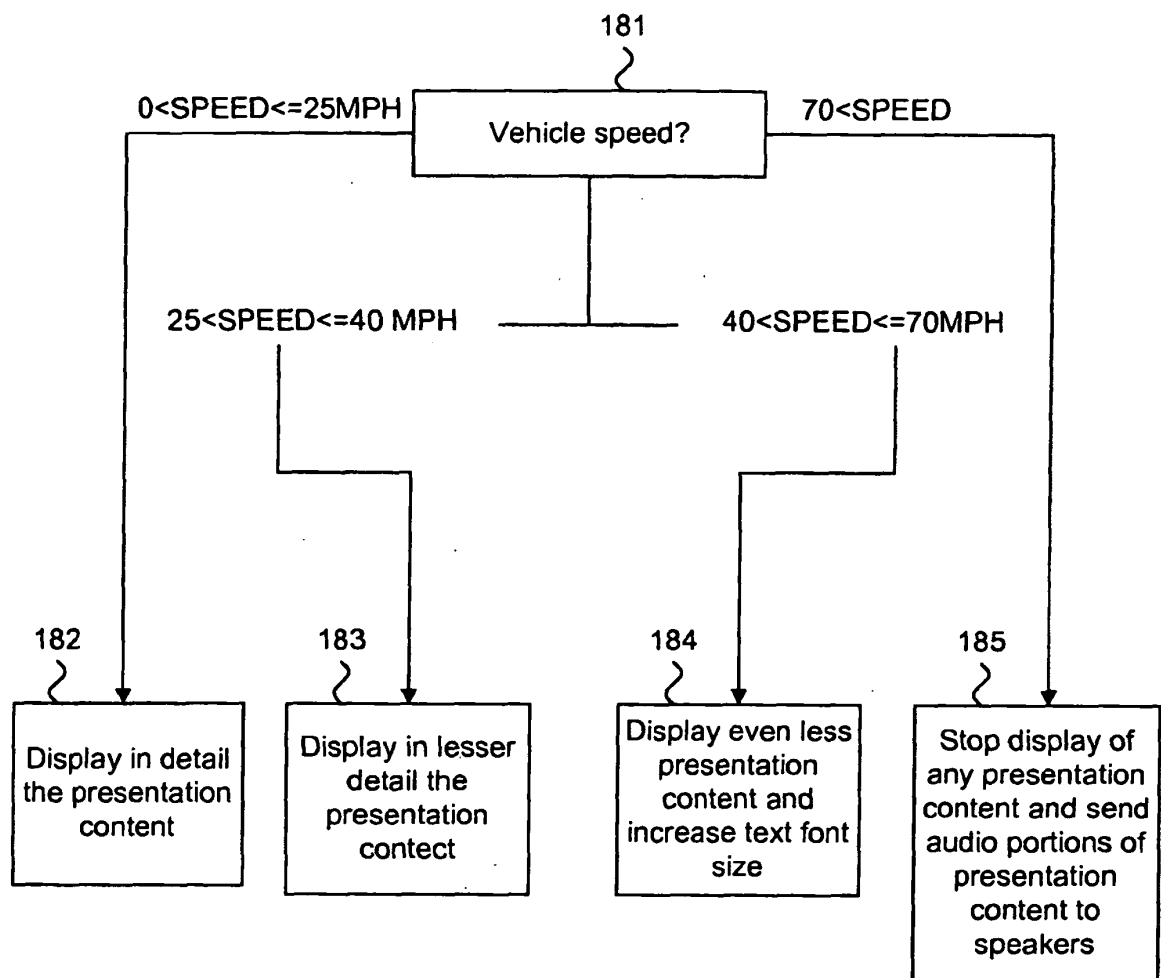
Fig. 16.180

Fig. 17.

190

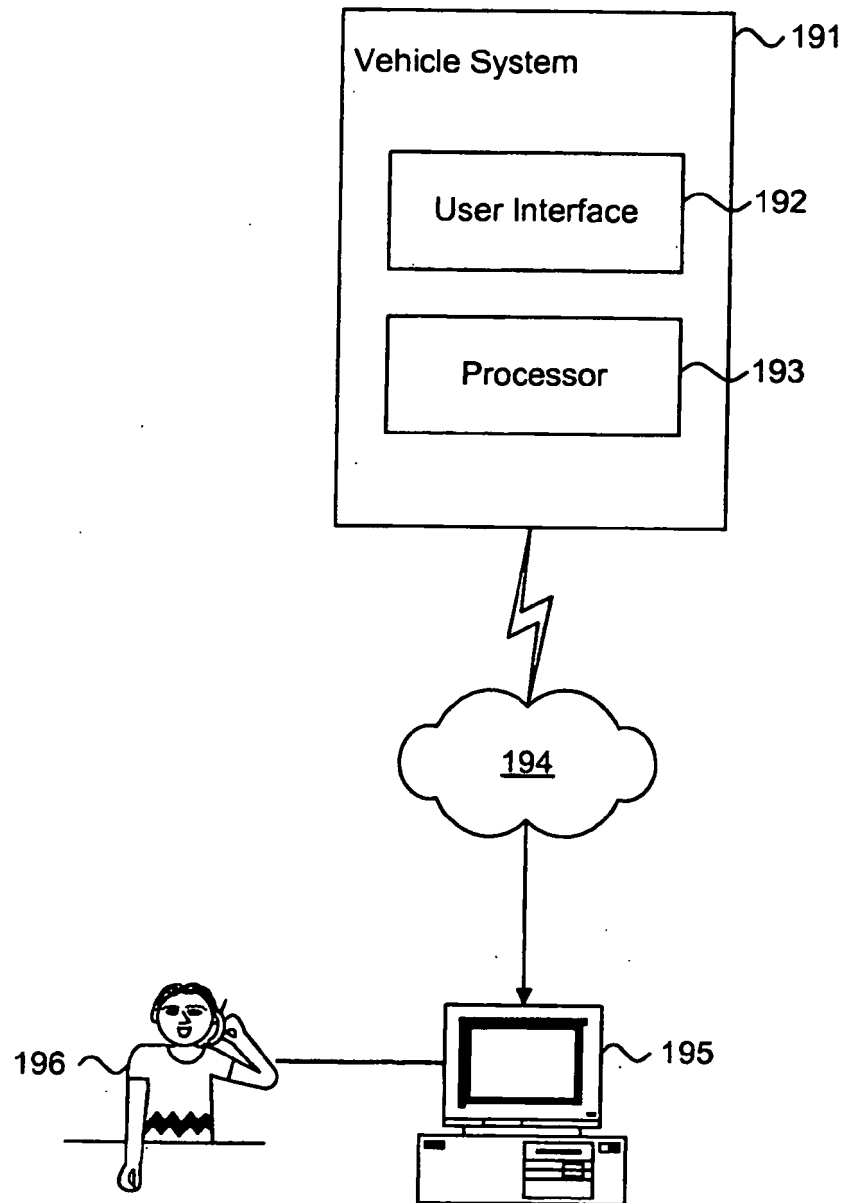
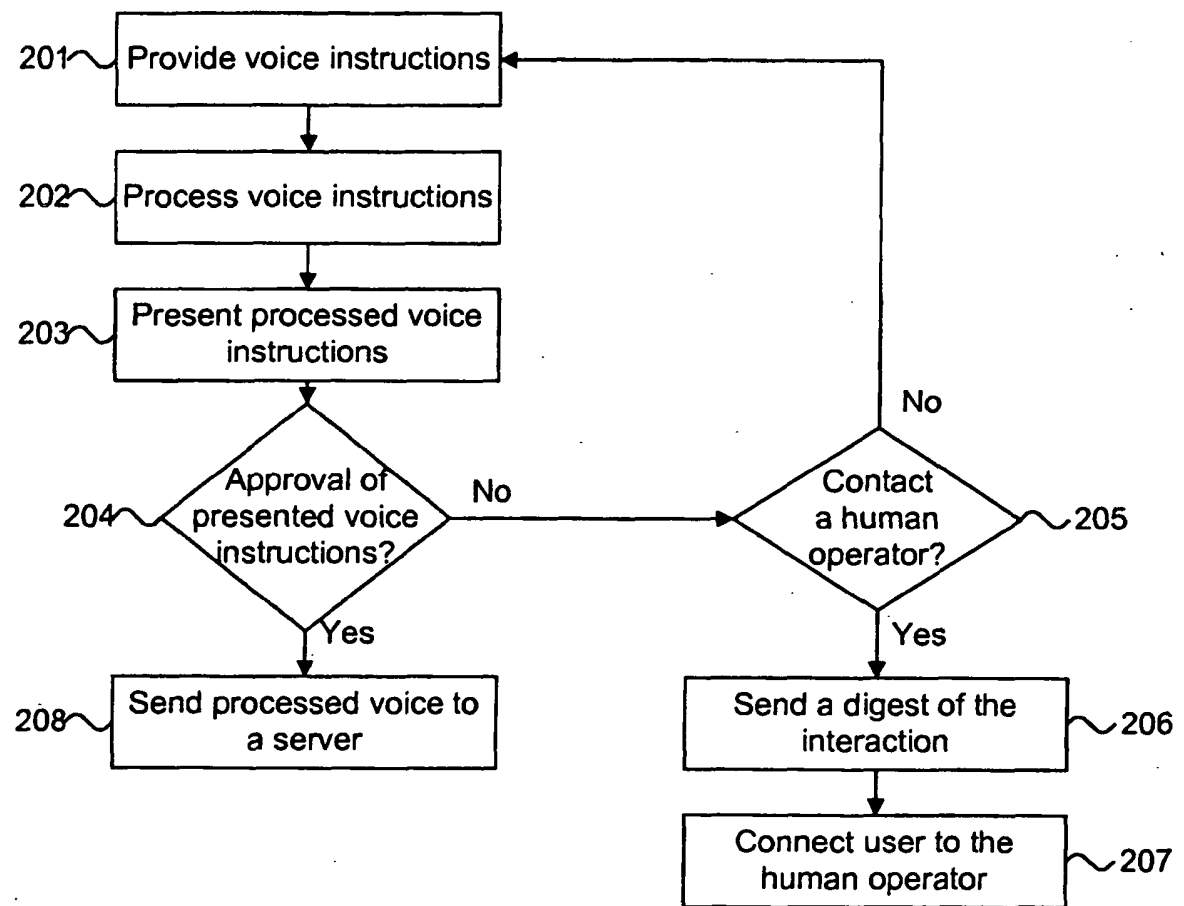


Fig. 18.200

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 6778660 B, Fromm [0007]
- US 2005147090 A1 [0010]
- US 2005213743 A1 [0011]